THE FRONTIERS COLLECTION

THE FRONTIERS COLLECTION

The books in this collection are devoted to challenging and open problems at the forefront of modern science, including related philosophical debates. In contrast to typical research monographs, however, they strive to present their topics in a manner accessible also to scientifically literate non-specialists wishing to gain insight into the deeper implications and fascinating questions involved. Taken as a whole, the series reflects the need for a fundamental and interdisciplinary approach to modern science. Furthermore, it is intended to encourage active scientists in all areas to ponder over important and perhaps controversial issues beyond their own speciality. Extending from quantum physics and relativity to entropy, consciousness and complex systems – the Frontiers Collection will inspire readers to push back the frontiers of their own knowledge.

Information and Its Role in Nature
By J. G. Roederer

Relativity and the Nature of Spacetime
By V. Petkov

Quo Vadis Quantum Mechanics?
Edited by A. C. Elitzur, S. Dolev, N. Kolenda

Life – As a Matter of Fat
The Emerging Science of Lipidomics
By O. G. Mouritsen

Quantum–Classical Analogies
By D. Dragoman and M. Dragoman

Knowledge and the World
Challenges Beyond the Science Wars
Edited by M. Carrier, J. Roggenhofer, G. Küppers, P. Blanchard

Quantum–Classical Correspondence
By A. O. Bolivar

Mind, Matter and Quantum Mechanics
By H. Stapp

Quantum Mechanics and Gravity
By M. Sachs

Extreme Events in Nature and Society
Edited by S. Albeverio, V. Jentsch, H. Kantz

The Thermodynamic Machinery of Life
By M. Kurzynski

The Emerging Physics of Consciousness
Edited by J. A. Tuszynski

Weak Links
Stabilizers of Complex Systems from Proteins to Social Networks
By P. Csermely

Peter Csermely

WEAK LINKS

Stabilizers of Complex Systems from Proteins to Social Networks

With 52 Figures and 12 Tables

Springer

Prof. Dr. Peter Csermely
Semmelweis University
Department of Medical Chemistry
P.O.Box 260
1444 Budapest 8
Hungary
e-mail: csermely@puskin.sote.hu

Cover figure: The cover image shows a detail from "Shock-Induced Vorticity" by J.O. Langseth

Library of Congress Control Number: 2006921156

ISSN 1612-3018
ISBN-10 3-540-31151-3 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-31151-5 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springer.com

Typesetting by Stephen Lyle using a Springer TEX macro package
Production and final processing by LE-TEX Jelonek, Schmidt & Vöckler GbR, Leipzig
Cover design by KünkelLopka, Werbeagentur GmbH, Heidelberg

Printed on acid-free paper SPIN: 11368847 57/3100/YL - 5 4 3 2 1 0

To the memory of my parents

Preface

In 1990 I started to work with molecular chaperones as an ordinary biochemist. Chaperones are the proteins that form our cells' most ancient defense system. I found them fascinating molecules. They protect other proteins and, consequently, help our cells to survive. If we quarrel, if we are anxious, or just run our daily marathon to catch the morning bus, our proteins become damaged. And damaged proteins are sticky. They aggregate, which is toxic to the cell. Chaperones protect these damaged proteins against unspecific, unplanned aggregation, like their eponyms, the ladies at the grand ball, who would protect young girls against unspecific, unplanned aggregation with the boys at the ball. Chaperones are everywhere. They are needed for protein folding and refolding, for proteolysis and transport. Chaperones are highly conserved and form a part of the essential gene set (Koonin and Galperin, 2002). Without them, no life could be imagined on Earth.

Chaperones are truly altruistic. They help, wherever they can. But how do they help? This was my first question. For five years I tried almost everything an ordinary biochemist could do. I purified them,[1] I cut them into pieces, cooked them and soaked them in an arsenal of chemicals and radioisotopes. By the middle of the 1990s, I realized that chaperones are different. They stick. They bind to their target proteins, their modulator proteins, the cytoskeleton, the whole world. If chaperones glue the whole cell together, how can it change? How do cells divide and how do they move?

The secret is affinity. Chaperones make *low* affinity interactions with their partners. Now they bind it, now they don't. They are dynamic. For their omnipresent help, they form weak links which change often. What makes life easy for the cell is a headache for the researcher.

[1] Footnotes will refer to additional information which is not needed to understand the main text. Therefore the reader may skip them. The first footnote is about the word 'purification'. We biochemists use this word for the procedure whereby we cut all original contacts of the protein and extract it from natural conditions, hoping and believing that it will remain unchanged.

Most of the chaperone complexes change if you start to examine them. Now you see them, now you don't. Chaperones give the ordinary biochemist nightmares. It is better to change subject if you want to use your usual assays in a sensible manner. I did not change subject – so the subject changed me. In 1990, I was looking for a well-defined question, and instead, I found a whole world with an astonishing complexity.

In 1998 a seminal paper by Suzanne L. Rutherford and Susan Lindquist appeared in Nature. The take-home message was that one of the chaperones, Hsp90, helps developmental stability. If this chaperone works, almost all the *Drosophilas* look alike. Each of these fruit flies gazes at the world with two complex red eyes, uses six small legs to balance her fragile body, and has two wings, which buzz in unison. If the Hsp90 was damaged, the newly born fruit flies went crazy. Fortunately, not all ten thousand of them did get damaged. Most of the flies still gazed, balanced and buzzed alike. However, *some* of them (exactly 174) became frightening monsters. These poor flies did not have correct eyes, their wings got distorted, their legs were deformed and a number of other malformations also occurred. Indeed they looked pretty miserable. But, miserable in *different* ways. An astonishing diversity appeared, and what is more, this diversity was inheritable. This was due to a variety of preserved silent modifications in the genome of the *Drosophila* population. Normally, Hsp90 buffered these changes and stabilized the appearance of the fruit flies, the phenotype. When Hsp90 was inhibited or damaged, the buffer diminished, and a burst of diversity suddenly appeared.

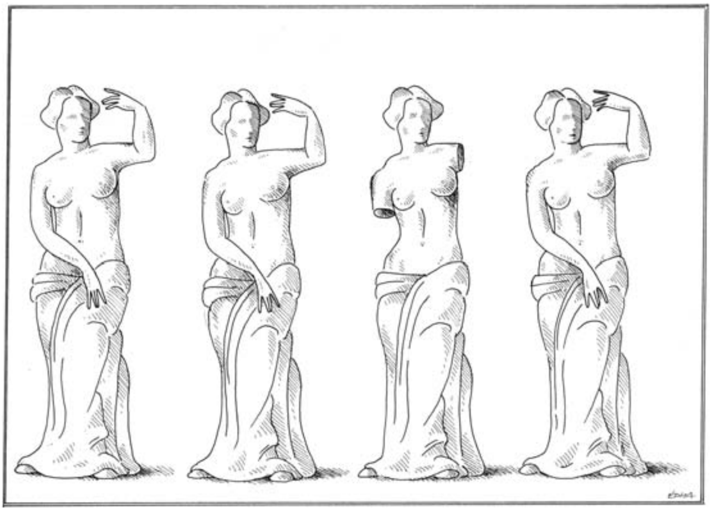
I got the feeling that something truly new had happened. Chaperones help the proteins around them. I could not figure out exactly how they do this, but at least I had an idea: chaperones bind to their target proteins and stabilize them or change their shape. But how can they stabilize a fruit fly, which is much bigger than them?

The explanation I offered myself was still quite standard, saying that chaperones repair mutant proteins which cannot exert their effects on the phenotype. Hsp90, the chaperone in the Rutherford and Lindquist (1998) experiments has hundreds of client proteins which always require its presence for their activation. Most of these clients participate in various steps of signal transduction. Let us suppose that the gene of one of these client proteins suffers a mutation, and that the mutation changes a critical amino acid and cripples the shape of the protein. Let us also suppose that Hsp90 is able to repair this damage, and that, if Hsp90 operates at full strength, the effect of the mutation is not seen. Finally, let us suppose that the client protein was critical in a signaling pathway of the morphological development. If Hsp90 is damaged, the mutation will impair the client. The missing client causes a collapse in morphological signaling and the fly will become a monster. The explanation seemed to be rather easy. (Well, it was easy for us, but not for the fly.)

There were disturbing signs though. Chaperones were not the only things that could hide the monsters within some of the normal *Drosophilas*, who seemingly gaze, balance and buzz just like their peers with a normal genome. There were numerous other proteins, which provided the same buffering, either in this or in other experimental systems (Aranda-Anzaldo and Dent, 2003; Gibson and Wagner, 2000; Scharloo, 1991; True and Lindquist, 2000). Moreover, in 2003 it was proposed that an astonishingly large number of proteins could regulate developmental stability (Bergman and Siegal, 2003). I became puzzled. I found chaperones fascinating. I loved them, and love always carries us to extremes. We see our beloved everywhere. Everything reminds us of her, she is everywhere. But wait a moment! Most of the proposed proteins had nothing to do with chaperones. Chaperones turn up here and there. But the whole cell cannot be a chaperone! The old explanation was clearly not adequate.

I think I am lucky. When one meets the unexpected, a fresh mind is needed, which finds an immense joy in each playful new thought. There are exceptional people, who have this even in their eighties. There are others, who are lucky enough to be stimulated by others. I started a project in 1996 giving research opportunities for high school students (`http://www.kutdiak.hu`). This movement changed the life of many students, and changed my life too. The students in my lab helped me to take a new look at the world. They were the seeds of the LINK group, who helped to write this book.

Let me put things together again. Rutherford and Lindquist (1998) showed that chaperones buffer the morphological diversity induced by the silent mutations of fruit flies. The inhibition of numerous other proteins can also lead to morphological diversity. However, there was something else here. Rutherford and Lindquist (1998) also demonstrated that stress induces a broader morphological diversity. In fact, the stress-induced, prolonged increase in morphological diversity was first shown by Schmalhausen and Waddington much earlier (Schmalhausen, 1949; Waddington, 1942; 1953; 1959). At first, I did not take much note here. It all seemed rather easy: perfect flies are alike, while stressed flies become damaged, *differentially* damaged. Diversity is revealed in the damage. At the molecular level, stress means more damaged proteins. Chaperones try to repair them, and so become occupied, which is just another form of inhibition.



Diversity is revealed through damage

In May 2003, I happened to read the review by Rao et al. (2002). This opened a new world to me. Stress not only induces morphological diversity, but also a thousand other types of diversity. Each bacterium normally swims towards its food. But not when stressed! Here, some of them got really distressed, and either swam in the opposite direction, moved round in circles, or just didn't go anywhere. Here we have diversity again. *Bacillus subtilis* responds to environmental stress with an arsenal of probabilistically invoked survival strategies. Stem cell differentiation or the appearance of various types of cancers can all be a source of similar diversity. Can all these forms of diversity be buffered by chaperones?

Putting this together, we are bound to ask: do we have here a large number of mysterious proteins which stabilize practically everything? This is the time when one goes for a vacation or asks around. As I was too excited to go for a vacation, I wrote to some of my best friends (I can imagine their faces as they stared at their laptops: "Peter went really crazy this time ...") – and got back some great ideas. Tamas Vicsek suggested that I read the recent book by Laszlo Barabasi on networks (*The Linked*, Barabasi, 2002). In parallel, I started to read *Investigations* by Stuart Kauffman (2000). These were the best books I had read for quite a while.

Then I had to sit down. Practically every complex system can be imagined as a network. Atoms form a network making macromolecules. Proteins form a network making cells. Cells form a network making organs and bodies. We form a network making our societies, and so on. Most of these networks are a result of self-organization. In fact, self-organization seems to be an inherent property of matter in our Universe. The resulting networks have a lot of common features, from their topology to their dynamism.

These systems are far too similar. The protein net, where chaperones work, should behave in the same way as every other network. *All* networks must have a component which stabilizes them, like chaperones and the mysterious proteins stabilizing the cells. But what is the common feature of all these elements? Why do they stabilize the whole network? At the beginning I had only one idea, and even this was negative. The common feature *cannot* be anything related to chaperone function. Chaperones protect other proteins by helping to refold them. People cannot protect their friends by helping to refold them! A more general approach is needed here.

Although I did not know it, the solution was already in my hands. Chaperones should give us a clue. Which of their features can be generalized to *all* networks? Chaperones stick. They make links to a number of other proteins. They are hubs. Do hubs stabilize their networks? Well, not really. Hubs are needed to *form* their networks. If we attack hubs, the network collapses (Albert et al., 2000). When we attack chaperones, the network, e.g., the cell survives. It becomes destabilized, but survives. Another chaperone feature must be more important.

What else? Affinity! Yes, affinity. Here was an idea: *The components which stabilize the various systems must all have weak links to the others*. It is not the component that counts, but the type of link it builds to the others. By the end of 2003, the basic idea of this book was born: weak links stabilize all complex systems. Weak links give us a universal key to understanding network diversity and stability, and they are the major actors in this book.

Months of tedious, systematic reading followed. I read dozens of books, collected approximately 600 Mbytes of pdf files, which made a pile of printed hard copies three meters high. I realized that my 'new' idea (weak links stabilize all complex systems) has been an obvious feature in the social sciences for decades (Granovetter, 1973). The same idea had been proven in ecosystems in 1998 (Berlow, 1999; McCann et al., 1998). As I browsed page after page, many other examples appeared, and they will be detailed in the following chapters. This made me rather confident that I had found something genuinely important and general. Interestingly, many authors (like Mark Buchanan in his book, *Nexus*, 2003) had come to the conclusion that weak links stabilize complex systems in their own discipline, but none of them had generalized it to all networks.[2] It seems that it was the chaperones, which stick but form only *weak* links, that had made the important link here.

While I was reading one book after the next and paper after paper, I got more and more surprises:

- I realized that each of the disciplines has a completely different vocabulary for the very same message. (Appendix B is a glossary intended to guide the reader through this jungle of terminology.)
- It was a frightening moment when the LINK group realized that we had completely run out of words and had no way of talking about something so truly simple and beautiful. But let me reassure you: the book is not full of newly constructed pseudo-words. We always managed to get around the problem ourselves and find a novel use for some existing word. However, on many occasions, it took us some time. When one has to use words in a completely different context, one's mind seldom obeys at first.
- The readings gave me a great and sincere respect for the social sciences. In network studies they are a whole lifetime ahead! Jacob Moreno started network studies on friendship patterns and Alfred Lotka published his famous law on scientific productivity in 1926, when my father was born. Anatol Rapoport stressed the general importance of the topology of friendship networks in 1957, one year before my own birth (Newman, 2003).

The book is structured as follows. In Chap. 1, I describe the beginnings of the weak link concept, the Granovetter study, and define weak links. Chapters 2 and 3 summarize the description and dynamics of networks. Chapter 4 introduces the concept of weak links as universal stabilizers, while Chaps. 5 through 11 invite the reader on a journey through Netland, presenting a ladder of exciting examples starting from macromolecules and ending at our own planet. Finally, Chap. 12 summarizes and reformulates the stabilizing role of weak links, bridging it with the

[2]The following remark by Siljak (1978) counts as another predecessor of these thoughts: "A dynamic system composed of interconnected subsystems is reliable if all subsystems are self-sufficient and [...] the magnitude of the interactions does not exceed a certain limiting value." This statement may be regarded as a forerunner of the main thesis of this book, but Siljak's stability criterion can be formulated much more easily using the concept of networks as it is presented here.

concept of stability landscapes and game theory. If you dislike physics or the biochemistry of small molecules, feel free to start your journey in Netland at Chaps. 7 or 8, which describe the networks of our own body and our societies.

When the first draft of the book was finished in January 2004, I realized that I had probably filled a niche. According to Newman (2003): "Studies of the effects of structure on system behavior are still in their infancy." Cross-disciplinary thinking on network properties is also largely non-existent. The moral is that we should use this enormous resource more often, always examining what we have proved in one of the disciplines when it is transferred to all the others. I have done my best. However, I am aware that analogies provide a very fruitful but extremely dangerous field. Therefore I will separate the analogies from the established facts by quoting the original source of information after each fact and by putting most of the analogies into a box in the following manner:

Caution! Hypothesis! As you proceed in the book, wild ideas will appear along the way. I expect a good deal of red ink from the referees: "Speculative!" But I have an excuse: I have *marked* all these hypotheses using one, two or three of the smiley figures on the left. The figure has big hands as a reference to the great Hungarian magician, Rodolfo, who always said: "Watch my hands! Caution! I am cheating!" One smiley means that I do not have enough evidence to formulate the statement as unequivocal truth. Two smileys warn you that, though the statement is logical, its background is largely missing. And three smileys? Well, three smileys will make you either smile or run to the phone to call the doctors in white coats. Three smileys are mostly fiction, rather than science. So why did I put them in this book? They have dared to appear here because they constitute fascinating, mind-boggling ideas. Smiley comments will always be exciting, but I am not quite sure that they will turn out to be true.

Additional information. Those parts where you find the wise head on the left will most probably turn out to be true tomorrow, and even the day after tomorrow, but they are details that will not necessarily interest all readers. Start reading, but if you do not find it interesting, skip it.

Important questions. When you begin to study a new territory, you always have more questions than answers. (In fact, a good scientist *always* has more questions than answers.) So we wondered why we should keep these questions to ourselves, and we decided to share. If you have a good idea for an answer, we would be more than pleased to read and discuss it. Join the LINKs! The email address is at the end of this Preface.

My master's voice: Spite. Sometimes you will see a remark in the text like: *"Peter, you made the typography of this book rather confusing for me. First of all, I cannot read your small letters in the remarks. Moreover, the font you selected for me is the ugliest one I have ever seen."* Spite! Welcome! Spite is my best friend. When you try to write a book, your best friend is the most critical person around you. I am lucky enough to have quite a few such fierce critics among my students.

Some of the sentences in the above remarks were written in the plural. What has happened? Does the author think he has found such a good idea that he may start to speak in the plural, as if to say: "We, the founders of this new science, declare ..."? Not so! The more 'we' know, the more humble 'we' grow. 'We' refers to the members of the LINK group. The LINKs are young people (at least in mind!), who work in different institutions but are strongly linked to each other by their love of weak links and decided to form a virtual lab. Members of the LINK group helped to shape this book. They sent great ideas to each other by emails, by SMS messages and even on slips of paper. Questions arose sometimes during the day, sometimes during the night. The LINKs attacked the sloppy sentences of this text and tried to make the content of the book more understandable. We all hope that this joint effort has brought at least a little improvement. If not, please send all your comments to us. Here are some of the key people in the LINK group:

Péter Csermely left the János Apáczai Csere high school in Budapest, Hungary in 1976. He won several awards in national and international chemistry contests. He is currently professor of biochemistry at the Semmelweis University in Budapest, Hungary and a fellow of Ashoka International. He has published nine books and almost two hundred research papers. He started a project in 1996 which provides research opportunities for more than seven thousand high school students in the best research teams (`www.kutdiak.hu`).

István Kovács left the János Berze Nagy high school in Gyöngyös, Hungary in 2003. He won awards in more than a dozen national physics and mathematics contests. He is currently a physics undergraduate at the Eötvös University of Sciences in Budapest, Hungary. He published his first scientific paper at the age of 19.

Balázs Papp left the high school of the Debrecen University, Hungary in 1996. He received his MSc degree in genetics at the Debrecen University and his PhD from the Eötvös University in Hungary. He is currently a postdoctoral fellow of the University of Manchester. He has published four papers in Nature, one in Nature Genetics, won several honors and awards including two Marie Curie Fellowships and a Pro Scientia Medal.

Csaba Pál left the István Dobó high school in Eger, Hungary in 1993. He received his MSc and PhD degrees from the Eötvös University in Hungary. He is in the Theoretical Biology Group of the Hungarian Academy of Sciences. He has published four papers in Nature, three in Nature Genetics, one in Science, as well as six papers in various Trends journals. He was a Royal Society Postdoctoral Fellow and won the Talentum Award of the Hungarian Academy of Sciences in 2005.

Máté Szalay left the László Lovassy high school in Veszprém, Hungary in 2003. He was the recipient of the 2003 Junior Bolyai Award of computer science. Between 2003 and 2005 he was the president and later the managing president of the Hungarian Research Student Association (`www.kutdiak.hu`). He is currently a computer science undergraduate at the Technical University in Budapest, Hungary.

An interdisciplinary subject is always dangerous. One cannot know, and cannot even understand everything. In spite of this, writing about weak links must not mean weak writing. I owe a lot of thanks to the eminent scientists of various disciplines, my friends, who read the summary of this book or its chapters. I am thankful for their comments, ideas and encouragement. The help of Luigi Agnati, Eszter Babarczy, László A. Barabási, Attila Becskei, Eric L. Berlow, Gustav Born, Zoltán Borsodi, Geoffrey Burnstock, György Buzsáki, Vilmos

Csányi, Ken Dill, Gerald M. Edelman, András Falus, Viktor Gaál, Balázs Gulyás, Mária Herskovits, Gergely Hojdák, Roland Iványi-Nagy, Gáspár Jékely, Ferenc Jordán, Márton Kanász-Nagy, Katalin Kapitány, Mária Kopp, Steve LeComber, Leon Lederman, Susan Lindquist, László Mérő, Ágoston Mihalik, István Molnár, Viktor Müller, Zoltan N. Oltvai, Kleopatra Ormos, Bálint Pató, Csaba Pléh, Zoltán Prohászka, Ricard V. Solé, Csaba Sőti, Attila Steták, Steven H. Strogatz, András Szabó, Péter Száraz, Gábor Szegvári, Attila Vértes, Tamás Vicsek, Denise Wolf and Peter Wolynes is gratefully acknowledged.

Without the encouragement of Tamás Vicsek, the help of our librarian, Csilla Szabó, and last but not least the editorial team, Claus Ascheron, Adelheid Duhm, James Fuite, Angela Lahee, Stephen Lyle and Jack Tuszynski, this book could not have been written. I would like to thank to all members of my family and my colleagues for their understanding during the writing process. Finally, let me introduce Édua Szűcs, to whom I am extremely thankful for the excellent artwork in the book.

Édua Szűcs left the Miklós Radnóti high school in Szeged, Hungary in 1977. She received her MSc from the Szeged University, Hungary. Starting her independent art work as a cartoonist in 1986, she has had more than thirty exhibitions in Hungary and abroad. Her published works include Edua cartoons (1997), Edua cartoons 2 (2001), and illustrations for several books. Awards: Szféra special awards (1996, 1999); Karikatórium special award (1997); Foundation for Hungarian Culture Award (1998); Women for the European Union, first prize (2003).

At the end of this preface, let me invite you once again to send us comments and questions. The LINK group can be reached at the following address and website:

`weaklink@puskin.sote.hu`   `www.weaklinks.sote.hu`

Budapest, Hungary *Peter Csermely*
September 2005

Contents

1 A Principle is Born: The Granovetter Study 1

2 Why Do We Like Networks? 5
2.1 Small-Worldness 7
2.2 Scale-Freeness 13
2.2.1 Scale-Free Degree Distribution of Networks 14
2.2.2 Underlying Reason for Scale-Freeness: Self-Organisation 18
2.2.3 Scale-Free Distribution in Time: Probabilities 24
2.2.4 Scale-Free Survival Strategies: Levy Flights 27
2.2.5 Scale-Free Pleasures 29
2.3 Nestedness 32
2.4 Weak-Linkness 43

3 Network Stability 47
3.1 Perturbations. Good and Bad Noise 47
3.2 Life as a Relaxation Phenomenon: Dissipate Locally, Connect Globally 56
3.2.1 Confined Relaxation with Global Connection 56
3.2.2 Self-Organised Criticality 59
3.3 Network Failures 68
3.4 Topological Phase Transitions of Networks 74
3.5 Nestedness and Stability: Sync 79
3.6 How Can We Stabilize Networks? Engineers or Tinkerers 90

4 Weak Links as Stabilizers of Complex Systems 95
4.1 An Emerging Synthesis: Weak Links Stabilize Complex Systems 95
4.2 Weak Links: A Starting Definition 100
4.3 Stability: A Starting Definition 103
4.4 Complex Systems 106
4.5 Weak Links and System Degeneracy 108

5 Atoms, Molecules and Macromolecules 111
5.1 Protein Folding Problems 111
5.2 Energy Landscapes 115
5.3 Weak Bonds in Protein and RNA Folding 119

6 Weak Links and Cellular Stability 125
6.1 Cellular Networks 125
6.2 Stability of the Cellular Net 128
6.3 Stress, Diversity and Jumps in Evolution 139
6.4 Cancer, Disease and Aging 150
6.4.1 Cancer 150
6.4.2 Disease 151
6.4.3 Aging 152

7 Weak Links and the Stability of Organisms 157
7.1 Immunological Networks 157
7.2 Transport Systems 162
7.3 Muscle Net 163
7.4 The Neuro-Glial Network 166
7.5 Psycho Net 170

8 Social Nets 181
8.1 Animal Communities 181
8.2 A Novel Explanation of the Menopause 184
8.3 Stability of Human Societies 186
8.4 Firms and Human Organisations 201
8.5 Dark Networks and Terror Nets 207
8.6 Pseudo-Grooming 209

9 Networks of Human Culture 219
9.1 The Language Net 219
9.2 Novels, Plays, and Films as Networks 224
9.3 Our Engineered Space 231
9.4 Software Nets 238
9.5 Engineers and Tinkerers: An Emerging Synthesis 240

10 The Global Web 243
10.1 The World Trade Web 243
10.2 Turning Points in History 247
10.3 Weak Links: A Part of Social Capital 258

11 The Ecoweb 263
11.1 Weak Links and the Stability of Ecosystems 263
11.2 Omnivory 267
11.3 The Weak Links of Gaia 270

12 Conclusions and Perspectives 275
12.1 The Unity of the Weakly-Linked World: A Summary . . 275
12.2 Revisiting the Definitions: A Synthesis 281
12.3 Prospects and Extensions 304
12.4 Weak Links and Our Lives 307

A Useful Links 313

B Glossary 317

References 329

Index 377

1 A Principle is Born: The Granovetter Study

In the late 1960s students had a rather revolutionary life in universities. In the midst of all this, Mark Granovetter, a PhD student at Harvard University, set himself to figure out how people find their jobs. He interviewed about a hundred people and sent out another 200 questionnaires in the Boston area.

The summary of his first results showed that more than half of the people found their jobs through personal contacts. We instinctively agree with these results. We may browse newspapers or web pages for a new job, but the real hints often come from our best friends. Or do they? In fact, this is not quite true. The really surprising result of the study was that, in most cases, the informants were not particularly close to the job seeker. They rarely spoke to each other, and they saw each other only seldom.

Why was this surprising? Granovetter had good reasons for thinking that strong links would be more useful for finding a new job. Close friends will give all their information to the job seeker and will mobi-

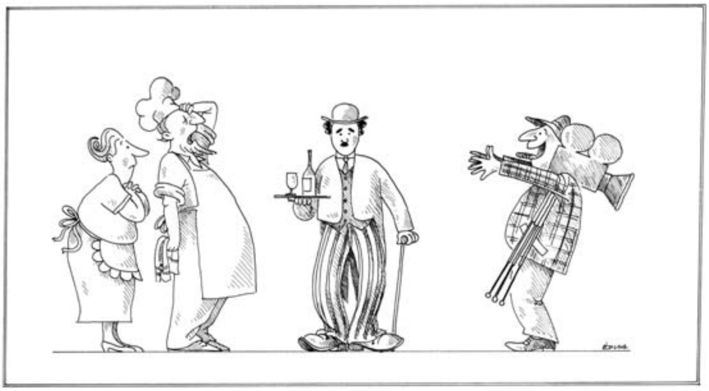


Fig. 1.1. In most cases, the best informants were not particularly close to the job seeker

lize all their contacts to help. Moreover, they meet the job seeker more often, and know more about her skills and preferences. And yet weak contacts still proved to be more useful. Were close friends biased? Did they overestimate the abilities of the job seeker?

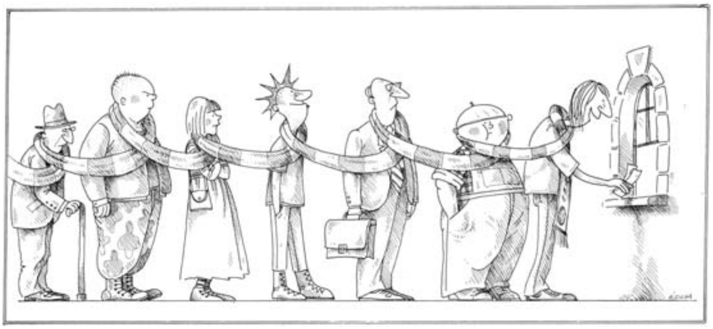


Fig. 1.2. Weak ties play a role in effecting social cohesion

Granovetter was puzzled and started to analyze earlier data. He considered an earlier hysteria incident, where more and more workers in a textile plant in the deep south of the USA were claiming bites from a mysterious and non-existent 'insect', until eventually the plant had to be closed (Kerckhoff et al., 1965). Although the rumor starters were isolated people, they had numerous weak links in the community. In Granovetter's meta-analysis, weak links also proved to be useful in the famous Milgram experiment (Milgram, 1967; Korte and Milgram, 1970). In this example, people were instructed to send a letter to an unknown person[1] in the USA by asking the help of persons they knew on a first-name basis. If the starter was white, and the target was an Afro-American, the 'chain of friends' worked efficiently only if the critical point, where the chain of white friends was switched to a chain of black friends, was a weak link. Finally, Granovetter showed that the friendship network of Rapoport and Horvath (1961) was best covered if one used weak links to search for the acquaintances of the acquaintances of a given person. In contrast, the 'best-friend' networks did not cover the whole community. It seemed to be a general result that weak links are more useful for information searches than strong ones.

[1]In the Milgram experiment only the postal address of the 'unknown person' (the target) was not revealed to the starter, and she did not know the endpoint personally. However, the starter did know the name and a few personal features of the target, e.g., the target is Rebecca Smith, a catholic Latin teacher in Cleveland, who is a chess champion.

Granovetter went further. He analyzed social networks in a general context, and observed that weak links also link network modules, a concept confirmed in many later studies. Finally, he came to the conclusion: "Weak ties play a role in effecting social cohesion." He published his findings under the title *The Strength of Weak Ties* (Granovetter, 1973). A principle was born. However, more than a quarter of a century was to pass before we started to learn that weak links not only connect, but also stabilize all complex systems. And now, we have reached

THE END.

Indeed we are already at the end. You have now heard the central statement of the book: weak links stabilize all complex systems. I described Mark Granovetter's landmark paper introducing this idea more than 30 years ago. I indicated the path leading to the generalization of this idea in the Preface. What more is there to say? *"How can you ask such a question? You have not even defined what you mean by 'weak links'?"* Thanks a lot, Spite, for the reminder. I will try to give a starting definition now, but if you would like to have a more complete version, please go ahead and check Sect. 4.2.

Weak links are links between network elements, which connect them with a low intensity. Weak links may also connect network elements with a higher intensity, but in this case they are only transient. I will show later that, in real networks, we have a continuous spectrum of link strengths starting with a few strong links and ending with more and more links, which become weaker and weaker. In most cases, it is rather difficult to cut the continuously changing strength parameter somewhere and say: up to here, all the links were strong, but from this point on, we shall say that they are weak. Consequently, in this book I will use the functional definition[2] of weak links given by Berlow (1999).

Definition of Weak Links. A link is defined as weak when its addition or removal does not change the mean value of a target measure in a statistically discernible way.

[2]It is a question of future research how much these 'functional weak links' overlap with the weak links, which are weak due to their low affinity or intensity.

The target measure here is usually an emergent property[3] of the whole network, or a response the network gives to a certain stimulus. The mean value of the target measure is changed if a strong link is deleted from or added to the network.

Will we lose weak links in the future? Please note that, in this functional definition, the discrimination between strong and weak links depends on the desired or available accuracy of our measurements. If the mean value is measured a hundred times more accurately, the 'statistically discernible change' in the mean value will be achieved by changing much weaker links than in the case of a measurement that is a hundred times less accurate. *"Why are you writing this book then? Your weak links will have vanished in a few years, when my generation has learnt how to measure things more accurately than your generation can."* I have bad news for you, Spite. When your generation has learnt how to measure things more exactly than we can measure them now, you will certainly lose a number of weak links according to this definition, since you will have to reclassify them as strong links. However, with the extension of detection limits, you will be able to measure a thousand times more 'new' weak links instead, which are even weaker than the weakest links my generation could detect. At the end of the day, your generation will have to deal with far more weak links than we ever did. As a conclusion, the younger you are, the more important this book is for you.

Having learnt a starting definition of weak links, this book will show that hierarchical networks are governed by the same principles, from molecules to the whole Universe, and that weak links stabilize us in all these levels. To understand all this, we must first learn more about networks. So let us begin.

[3]For the explanation of the meaning of 'emergent property' and other unusual words in the text, please see the glossary in Appendix B.

2 Why Do We Like Networks?

Networks catch hold of you. They are enchanting and contagious. As a first 'proof' of these statements let me give you my own example. Just before starting to write this chapter, I sat on a train and watched a charming mother and her little daughter just opposite me. The baby fell asleep playing with her comforter. As I continued to watch, my mind went to work. What could be the periodicity of her suckling motions? Was it perhaps scale-free, showing sudden bursts of activity separated by longer and longer periods of stasis? Did it show self-organized criticality? Was this a punctuated equilibrium? My thoughts continued: What if I looked outside? Would I see fractals instead of trees and clouds? *"Let me interrupt you here. Why do you assume that we know what 'scale-free', 'self-organized criticality', 'punctuated equilibrium' and 'fractal' are supposed to mean? And anyway, what is a network?"* I am sorry, Spite. Whenever elements are connected with links, we may call them a network. Networks can be formed from atoms, molecules, cells, plants, firms, words, power stations, Internet routers, Web pages, countries, etc. Even your friends, Spite, form a network. The meaning of the other words will be explained later. If you are curious to understand them now, turn to the glossary in Appendix B, at the end of the book.

Returning to the popularity of networks, I am not the only one who has found this field fascinating. Figure 2.1 shows the number of network-related scientific publications in MEDLINE.[1] The arrow points to the publication date of two important network discoveries, the demonstration of the generality of the small-world phenomenon (Watts and Strogatz, 1998) and scale-free behavior (Barabasi and Albert, 1999). Obviously, these data may just reveal a coincidence. They do not directly prove the profound effects of these important discoveries in the network approach. However, Fig. 2.2 shows the number of

[1]Data of in Fig. 2.1 should be treated with caution, since 'network' may also refer to a network of authors, for example. An additional non-specific effect arises from the fact that the number of annual publications covered by MEDLINE also increased over the period covered.

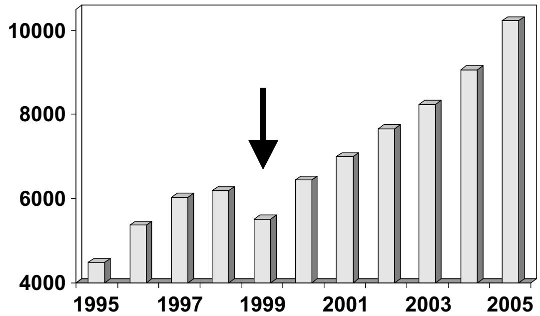


Fig. 2.1. Number of network-related publications in MEDLINE. The number of publications containing the words 'network' or 'networks' in their title or abstract was collected from MEDLINE (`www.pubmed.com`). The 2005 data is an extrapolation. The *arrow* shows the publication date of the two seminal network papers by Watts and Strogatz (1998) and Barabasi and Albert (1999)

annual citations of the above two papers in comparison with the average citations of three randomly selected papers from the same journals having a similar number of total citations. The citations of randomly selected papers peter off after 3 to 4 years. In contrast, citations of the two seminal network papers grow linearly, showing no tendency to decline in this period. No wonder, scientists seem to like networks. But how about the layperson? As a measure of public success, Laszlo Barabasi's book, *The Linked* was translated into 8 languages in the first two years of its existence.

Having these data to hand, I think we may be quite confident in saying that networks really catch hold of people. People do like networks. However, another question arises: Why exactly do people like networks? This chapter attempts to answer this question and uses the elements of the answer to introduce some important features of networks in general.

Small-worldness, scale-freeness, nestedness, weak-linkness: these are the titles of the following sections. All these words refer to properties which are general features of most networks around us, and this is why we have acquired a feeling for them. These properties of the networks we either contain or belong to inherently help us to understand the world around us, being basic, underlying elements of our cognition. Therefore small-worldness, scale-freeness, nestedness and weak-linkness not only mean the actual features of networks (being a small-world network, having scale-free distribution of various properties, containing other networks as its elements as well as belonging to higher order networks and having a large number of weak links,

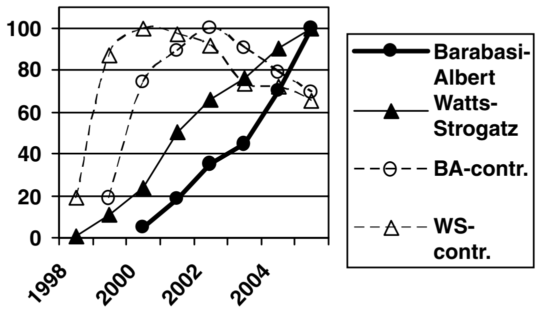


Fig. 2.2. Citations of seminal papers on networks. The numbers of citations for the Watts and Strogatz (1998) and Barabasi and Albert (1999) papers were collected from the Web of Science. Control values show the number of citations of three randomly selected papers from the same journals having a similar total number of citations. Data were normalized to the maximal number of yearly citations. 2005 data is an extrapolation

respectively), but also refer to the help these network properties give us. What is this help? Please continue, if you would like the answer.

2.1 Small-Worldness

Stanley Milgram did many famous experiments. In his small-world experiment he gave letters to starters, persons, who were asked to pass them to acquaintances known on a first-name basis in order to find an unknown, distant target (Milgram, 1967). Imagine that you have the task of sending a letter to the Reverend Lucas Brown, who lives in the capital of Myanmar, Yangon. It is rather easy. I need the address, ZIP code and a few stamps. But not this time! No address is known, and direct mailing is excluded. You may pass the letter only to one of your friends. The important message of Milgram's work, viz., "we live in a small world, and are only six steps apart from each other", became very popular. There is a good chance that your letter to the Reverend Brown will find its target by passing along a chain of around six friends.

A Hungarian prediction of small worlds from 1929. Tibor Braun (2004) quotes the following text from a story by the Hungarian writer, Frigyes Karinthy, in 1929: "To prove that nowadays the population of the Earth is in every aspect much more closely interconnected than it ever has been, one member of our gathering proposed a test. 'Let us pick at will

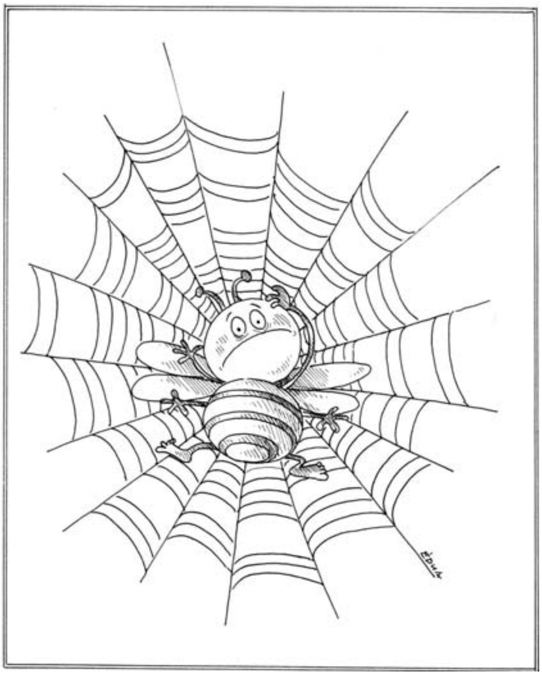


Fig. 2.3. Networks can really catch hold of you

any given existing person from among the one and a half billion inhabitants of the Earth, at any location.' Then our friend bet that he could establish via direct personal links a connection to that person through at most five other persons, one of them being his personal acquaintance. 'As people would say, look, you know X.Y. Please tell him to tell Z.V., who is his acquaintance, and so on.' 'OK,' said a listener, 'then take for example Zelma Lagerlöff' (Nobel Prize for Literature, 1909). Our friend placing the bet remarked that nothing could be easier. He thought for only two seconds. 'Right,' he said, 'so Zelma Lagerlöff, as a Nobel Laureate, obviously knew the Swedish king Gustav, since the king handed her the prize, as required by the ceremony. Gustav, as a passionate tennis player, who also participated at large international contests, evidently played with Kehrling [Béla Kehrling (1891–1937), Hungarian tennis champion and winner at the Göteborg Olympics 1924], whom he knew well and respected.' 'Myself,' our friend said (he was also a good tennis player), 'I know Kehrling directly.' Here was the chain, and only two links were needed out of the stated maximum of five." The amazing foresight of Karinthy (1929) predicting that we are approximately five steps apart from each other on the global scale was proved decades later by Milgram (1967) and Dodds et al. (2003a).

When I gave a lecture on networks to illustrate the smallness of the small world we live in, I asked my audience how many steps they thought they were from the President of the United States. Some of them guessed around a hundred, others were better informed and said: Six! Then I surprised them with the exact number: Three. *"How come? Did you know that someone's parent attended the same school as the President?"* No, Spite, I knew only my own connections. I happen to know the President of my country, Hungary, who did meet the President of the USA. Since the students knew me, this is exactly three steps for them. Our world is really small. However, there is another message from the Milgram experiment: not only do short paths exist between distant network members, but ordinary people are very good at finding them too (Newman, 2003b). How would you kick off your letter to the Reverend Brown in Yangon? *"I happen to have a friend who moved to Kuala Lumpur a year ago. If I recall my geography lessons, it is not far from Yangon. I would ask her to look around. She certainly knows many more people in the region than I do."* Excellent, Spite. If she happens to know a priest in Kuala Lumpur or Yangon, you might even complete the chain in three steps instead of six.

Why was Milgram lucky? Examining the original numbers, I have to conclude that in spite of the seemingly easy navigation shown above, Milgram was lucky. The final conclusion was based on only 18 letters which actually reached the single target of the Milgram experiment in Boston out of the 96 starters at distant locations in Nebraska. In other studies the success rate was even lower (Kleinfield, 2002). It was often hard to define what was causing the numerous drop-outs. However, a later study (Dodds et al., 2003a), using tens of thousands of emails had the same conclusion: we are about six steps apart even in different parts of the world. Experimenters of robust phenomena are lucky. Their instincts often find the right solution even when the actual proof is shaky. However, if you are a young investigator, let me ask you *not* to rely on this. Unsuccessful examples always outnumber the few serendipity stories. We do not hear about the failures: most of them never get published. Moreover, our publication habits mean that we mention only the final success stories and not the very important and frustrating path we had to follow to reach them.

The number of dimensions in our brain. How do we select a direction to send our letter towards the unknown target? In fact, we try to get a match between the character of our acquaintances and the known properties of the target. For this matching task, we categorize our friends into social dimensions, as has been shown in the model by Watts et al. (2002). Spite started the letter to the Reverend Brown by finding a friend in Kuala Lumpur. This was a wise tactic, since geographic information is sufficient to perform global routing in a significant fraction of cases (Liben-Nowell et al., 2005). However, I have no friends in the region, and therefore I would probably double check the priests in my circle of friends instead. The number of social dimensions screened for a search is around 5 to 6 (Dodds et al., 2003a; Killworth and Bernard, 1978). This number is actually quite close to our average cognitive dimensionality, which is measured as the number of persons whose intentions towards each other I may still follow (Dunbar, 2005). What should we do if we want to be even more sophisticated? Should we use more social dimensions? I might have bad news here. The dimensionality of our neural network may prevent this. Even if our world grows hopelessly complex, we will still restrict ourselves to half a dozen, or even fewer, social dimensions to describe it, or start to develop more complex neural nets in our brain. Watch out for contact gurus! The evolution of superbrains may actually be happening around you now as you read!

Our world is a small world. However, it is not only the expanding circles of friends, the social net, which is a small world. Many other networks, like power nets, the networks of neural cells, etc., are also small-world networks. We live in small-worldness. *"This is obvious. If I take a hundred people and everyone knows everyone, their world – let me call it Spiteland for short – is really small. Not six steps, but one, separating any one of them from any other."* Spite, I appreciate your logic, but real life is not Spiteland. We cannot know everyone. Do you have six billion friends, increasing by dozens every second? I doubt it. However, you have hit upon a good point here. Small worlds are not only small in the sense that their members may reach each other easily. Random graphs, where the connections are made between the elements in a random fashion, are equally good at this (see Fig. 2.4). In small worlds, your neighbors also know each other. To use a scientific term, this 'my friend's friend is my friend' effect is called clustering. The clustering of small-world networks is high. These networks are lucky mixtures of 100% clustered regular lattices and highly-connected random graphs (see Fig. 2.4) (Watts and Strogatz, 1998). Small-worldness requires both a dense array of local contacts, which is reflected by the high clustering of small worlds, and a good enough number of long-range contacts. The simultaneous presence of both ensures that the small-world network becomes really small, providing easy conditions for finding any of its members. However, this cannot be achieved by extensive cross-linking

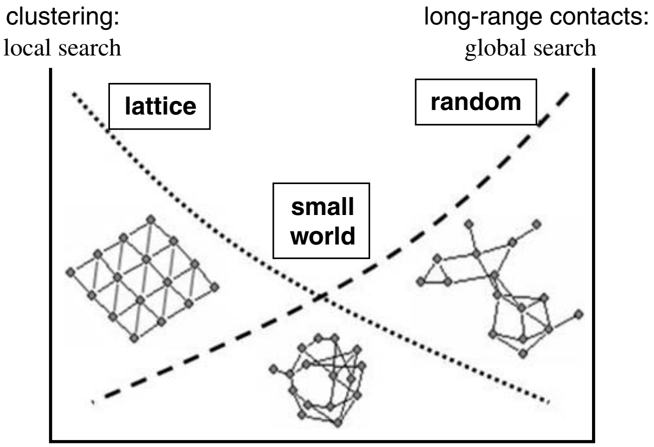


Fig. 2.4. The small-worldness of networks. The figure shows that small-world networks are in-between lattice-type networks and random graphs, having much longer range contacts than the former and much higher clustering than the latter. Note that the measures of both clustering and long-range contacts are purely illustrative

of the network, since building and maintaining links is costly. Natural small worlds are economical (Latora and Marchiori, 2003). In fact, small worlds are much more economical than either random networks or regular lattices.

Some worlds are not so small. Small-worldness depends on what we consider as a member of the network. As an example, the extent of the small-world status of metabolic networks may vary, if we include relatively simple molecules like water, ATP, use directed links, or restrict the network to conserved residues of participating molecules (Arita, 2004; Ma and Zeng, 2003).

How many friends do you need to send your message to anyone? In 2000, Jon Kleinberg published an interesting model for message transmission on a two-dimensional lattice, where lattice elements were linked with random shortcuts. The interesting result was that an optimal condition can be defined for the fastest search. If the shortcuts are neither fully random (where lots of short paths exist, but it is extremely time-consuming to find them), nor restricted to short-range contacts (where no short paths exist at all), an optimal condition can be found where the system transmits the messages most efficiently. Under these conditions, you have exactly the same number of friends in your neighborhood, in the rest of your city, in the rest of your country, in the rest of your continent and in the rest of the world. In other words, you only have to worry about how to send your message to someone in the right neighborhood. Once the message has reached the right region, the fine-tuned targeting will rely on the increasingly denser local contacts as the message homes in on the actual target. This makes the search highly efficient and the system behaves like a small world (Kleinberg, 2000). The Kleinberg condition can be reformulated: for an optimal two-dimensional search, if you go to a higher region (neighborhood, city, country, continent, world), the chances of finding a friend after a random selection become an order of magnitude smaller. Kleinberg's model behaved optimally if the number of connections was the same on all scales, i.e., it was scale-free. Scale-freeness is another important feature of our everyday networks besides small-worldness, and will be discussed in detail in the next section.

Small worlds are easy to navigate. Lattice-type connections with their high clustering ensure the success of the finely-tuned final steps of a target search. Long-range contacts ensure the success of initial steps zooming in on the region of final interest. "The key to generating the small-world phenomenon is the presence of a small fraction of [...] edges, which contact otherwise distant parts of the graph" (Watts, 1999).

Navigation in small-world networks is helped by weak links (Granovetter, 1973; Lin et al., 1978). In many networks such as social networks, most of the long-range links typical of small worlds are weak links (Onnela et al., 2005). As I mentioned above, Dodds et al. (2003a) repeated the Milgram (1967) experiment using more than 60 000 emails. They found that a successful social search was conducted primarily through intermediate-to-weak strength links and did not require highly connected hubs. Moreover, Skvoretz and Fararo (1989) showed that the more weak links there are in a population, the closer a randomly chosen starter is to all others.

Interestingly, groups with lower or higher socioeconomic status, as well as groups under stress, tend to use strong links instead of weak ones (Granovetter, 1983; Killworth and Bernard, 1978). As a possible consequence of this, people under stress and either on the top or at the bottom of society may belong to a more closed world than those living in relative rather than extreme prosperity.

Why do we like small-worldness? Why do we need it? Humans are cooperative animals (Ridley, 1998). Consequently, our brain has devel-