Emergence, Complexity and Computation ECC

Ali Sanayei
Ivan Zelinka
Otto E. Rössler *Editors*

ISCS 2013: Interdisciplinary Symposium on Complex Systems

Springer

Emergence, Complexity and Computation

Volume 8

About this Series

The Emergence, Complexity and Computation (ECC) series publishes new developments, advancements and selected topics in the fields of complexity, computation and emergence. The series focuses on all aspects of reality-based computation approaches from an interdisciplinary point of view especially from applied sciences, biology, physics, or Chemistry. It presents new ideas and interdisciplinary insight on the mutual intersection of subareas of computation, complexity and emergence and its impact and limits to any computing based on physical limits (thermodynamic and quantum limits, Bremermann's limit, Seth Lloyd limits…) as well as algorithmic limits (Gödel's proof and its impact on calculation, algorithmic complexity, the Chaitin's Omega number and Kolmogorov complexity, non-traditional calculations like Turing machine process and its consequences,…) and limitations arising in artificial intelligence field. The topics are (but not limited to) membrane computing, DNA computing, immune computing, quantum computing, swarm computing, analogic computing, chaos computing and computing on the edge of chaos, computational aspects of dynamics of complex systems (systems with self-organization, multiagent systems, cellular automata, artificial life,…), emergence of complex systems and its computational aspects, and agent based computation. The main aim of this series it to discuss the above mentioned topics from an interdisciplinary point of view and present new ideas coming from mutual intersection of classical as well as modern methods of computation. Within the scope of the series are monographs, lecture notes, selected contributions from specialized conferences and workshops, special contribution from international experts.

Ali Sanayei · Ivan Zelinka
Otto E. Rössler
Editors

ISCS 2013: Interdisciplinary Symposium on Complex Systems

Springer

Editors
Ali Sanayei
University of Tübingen
Tübingen
Germany

Ivan Zelinka
Department of Computer Science
VŠB-Technical University of Ostrava
Ostrava
Czech Republic

Otto E. Rössler
Institute of Physical and Theoretical Chemistry
University of Tübingen
Tübingen
Germany

ISSN 2194-7287 ISSN 2194-7295 (electronic)
ISBN 978-3-642-45437-0 ISBN 978-3-642-45438-7 (eBook)
DOI 10.1007/978-3-642-45438-7
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014931496

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The book you hold in your hands is the outcome of the "2013 Interdisciplinary Symposium on Complex Systems" held at the historical capital of Bohemia as a continuation of our series of symposia in the science of complex systems. Prague, one of the most beautiful European cities, has its own beautiful genius loci. Here, a great number of important discoveries were made and many important scientists spent fruitful and creative years to leave unforgettable traces. The perhaps most significant period was the time of Rudolf II who was a great supporter of the art and the science and attracted a great number of prominent minds to Prague. This trend would continue. Tycho Brahe, Niels Henrik Abel, Johannes Kepler, Bernard Bolzano, August Cauchy Christian Doppler, Ernst Mach, Albert Einstein and many others followed developing fundamental mathematical and physical theories or expanding them. Thus in the beginning of the seventeenth century, Kepler formulated here the first two of his three laws of planetary motion on the basis of Tycho Brahe's observations. In the nineteenth century, nowhere differentiable continuous functions (of a fractal character) were constructed here by Bolzano along with a treatise on infinite sets, titled 'Paradoxes of Infinity' (1851). Weierstrass would later publish a similar function in 1872. In 1842, Doppler as a professor of mathematics at the Technical University of Prague here first lectured about a physical effect to bear his name later. And the epoch-making physicist Albert Einstein—while being a chaired professor of theoretical physics at the German University of Prague—arrived at the decisive steps of his later finished theory of general relativity during the years 1911–1912. In Prague, also many famous philosophers and writers accomplished their works; for instance, playwright Karel Čapek coined the word "robot" in Prague ('robot' comes from the Czech word 'robota' which means 'forced labor').

We believe that the Prague genius loci will vibe fruitfully with the conference's topic. Especially since the paradigm itself—the "Science of Complex Systems"—does not have a unique meaning so far since there still does not exist a canonical definition of the word complexity given the ongoing presence of divergent points of view. In the face of this fact we hear the reader asking us: What is the reason that we push on with our series of symposia with its amalgamation of different fields and viewpoints which sometimes do not appear consonant with each other?

We would reply with a single word: “Creation”. This answer we take from Paul Dirac who was once asked by Michael Noakes—a portrait painter of the British royal family—“Can you put into layman’s terms what you are working on, Professor?” “Creation”, was the reply he got. When Noakes, deeply amazed by such a short profound answer, asked for more explanation, he heard: “Creation was one vast bang. Talk of a steady state is nonsense.” When he tried to throw-in that if nothing had existed beforehand, what was there to bang, he would meet with an elegant evasion: “That is not a meaningful question!”

In this spirit of an anticipated but not yet existing elegance, the reader—student or professional scientist—will in the following encounter four general categories of papers in the following: Fundamental and theoretical, systemic modeling, systemic networking, and applications of various kinds. All papers represent innovative ideas, philosophical overviews or state-of-the-art applications in miscellaneous fields. At this point we would like to express our gratitude to our participants, our program committee members, and all our friends at the Czech Technical University in Prague who helped us continue our symposia at a progressive rather than “steady state” rate. We also acknowledge our keynote speakers in this year’s symposium for sharing their results: Yaneer Bar-Yam, Florentino Borondo, Robert Devaney and Stefan Thurner.

As a token of appreciation we dedicate the book to Graham Farmelo who revived Dirac’s attitude in science—“Physical laws should have mathematical beauty” and “creation means not to get locked up in a steady state”—in his penetrating 2009 biography which opens up a new way to look at the world as a single complex system.

August 2013

Ali Sanayei
Ivan Zelinka
Otto E. Rössler

Contents

Part I
Complex Systems Science

The Complex Geometry of the Mandelbrot Set

Robert L. Devaney

Abstract In this paper, we give a brief overview of the geometry of the Mandelbrot set. We show how to distinguish each of the principal bulbs hanging off the main cardioid of this set by counting the spokes of the antennas attached to each bulb. We also use these antennas to attach a fraction to each such bulb, and this then indicates how these bulbs are arranged around the boundary of the main cardioid.

Keywords Mandelbrot set · Julia set · Fatou set · Complex dynamics · Fractal · Farey addition · Quadratic dynamics

The Mandelbrot set $\mathcal{M}$ is one of the most interesting and beautiful objects in all of mathematics. It is also one of the most intricate planar sets. Contrary to the fact that it is named after Benoit Mandelbrot, the father of fractal geometry, this set is the antithesis of a fractal in that almost every tiny area of the boundary has its own "identity." That is, using some tools from geometry, we can read off exactly where this boundary point is and, more importantly, exactly what the corresponding dynamical behavior is.

Amazingly, the Mandelbrot set arises as the parameter plane for the seemingly simple quadratic family $P_c(z) = z^2 + c$. See Fig. 1. This is a picture in the c-plane (the parameter plane) that describes the fate of the orbit of the only critical point for this family, namely 0. If the orbit of 0 does not tend to ∞, then the corresponding parameter c lies in $\mathcal{M}$ and we color this point black. If the orbit does escape to ∞, then c is not in $\mathcal{M}$ and we color c according to how quickly the orbit of 0 reaches the exterior of a large disk surrounding the origin (with red points escaping fastest, followed in order by orange, yellow, green, blue, and violet).

2000 MSC number: Primary 37F10; Secondary 37F45
This work was partially supported by grant #208780 from the Simons Foundation.

R. L. Devaney (✉)
Department of Mathematics, Boston University, 111 Cummington Mall,
Boston, MA 02215, USA
e-mail: bob@bu.edu

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_1,

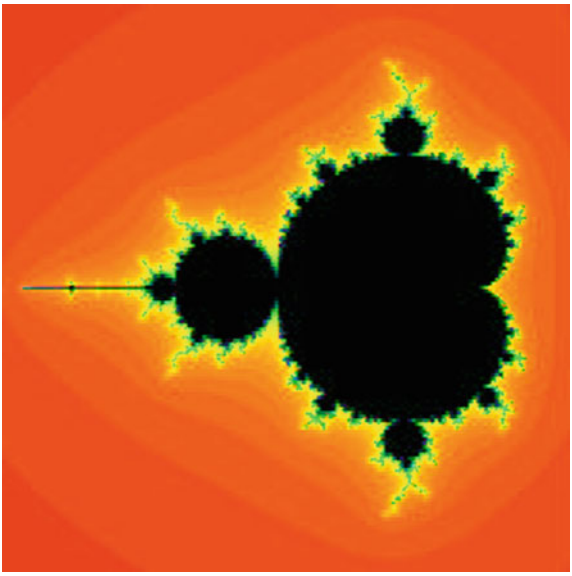


Fig. 1 The Mandelbrot set. Colored points are c-values for which the orbits of 0 escape to ∞; black points are c-values for which this does not happen. So the Mandelbrot set is the black region in this image

In complex dynamics, the object of central interest in the dynamical plane is the *Julia set*. For the family P_c, there is an open neighborhood of ∞ in the Riemann sphere consisting of points whose orbits tend to ∞. The set of all points whose orbits tend to ∞ is called the basin of ∞. Then the Julia set, denoted by $J(P_c)$ is the boundary of this basin. Given any point in the Julia set, then any open neighborhood of this point, no matter how small, is eventually mapped over the entire complex plane, minus at most one point. So the family of iterates of P_c on the Julia set is very chaotic.

The natural question is: Why are we interested in the fate of the orbit of the critical point? Well, in short, the critical orbit "knows it all" in complex dynamics. In particular, for the family P_c, if the orbit of 0 tends to ∞, then the Julia set of P_c is a Cantor set, i.e., a scatter of infinitely many points, or "fractal dust." If the orbit of 0 does not escape to ∞, then $J(P_c)$ is a connected set, i.e., just one piece.

The large black open regions (called *hyperbolic components*) visible in the Mandelbrot set are regions for which P_c has an attracting cycle of some given period. It is known that, if P_c has an attracting cycle, then the orbit of the critical point must tend to this cycle. Hence there can be at most one attracting cycle for a quadratic polynomial. For example, any c-value drawn from the central cardioid has an attracting fixed point. For c in the large open disk just to the left of this cardioid, P_c has an attracting 2-cycle. We therefore call this the period 2-bulb. And, for c in the northernmost and southernmost bulbs off the main cardioid, P_c has an attracting cycle of period 3, so these are the period 3-bulbs.

As c moves from one hyperbolic component to another, the map undergoes a bifurcation. The simplest part of this bifurcation is the fact that we move from having

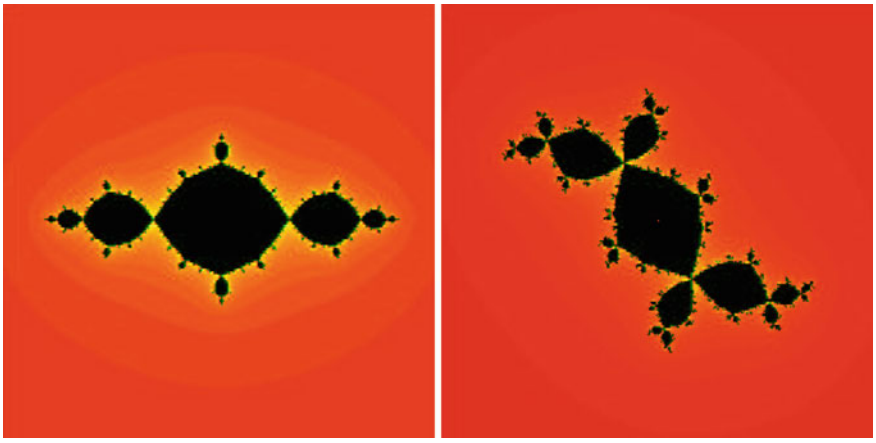


Fig. 2 The Julia sets for $z^2 - 1$ (the basilica) and $z^2 - 0.12 + .75i$ (the Douady rabbit). The filled Julia sets are the black regions, so the Julia sets here are the boundaries between the black and colored regions

an attracting cycle of some period when we are in one hyperbolic component to having an attracting cycle of some other period in the subsequent hyperbolic component. But, in fact, much more happens: the topology of the Julia sets changes dramatically. For example, if we move from the main cardioid to the period-2 bulb, the Julia set, which is just a simple closed curve when c is in the main cardioid, becomes a "basilica" when c is in the period 2-bulb. What happens is a repelling 2-cycle that lies in $J(P_c)$ when c is in the cardioid suddenly merges with the attracting fixed point and thereby makes it neutral when the parameter reaches the boundary of the cardioid. So two points in $J(P_c)$ become identified to one point. Meanwhile, infinitely many pairs of preimages of this point also become identified. This is what accounts for the infinitely many "pinch-points" visible in the basilica. Or, as we move from the main cardioid to the period 3-bulbs, a period 3-cycle becomes identified and the Julia set transforms into the "Douady rabbit". See Fig. 2.

A natural question is how do we understand how all of the bulbs and other smaller Mandelbrot sets are arranged in $\mathcal{M}$. Amazingly, if we zoom in to any portion of the boundary of the Mandelbrot set, it turns out that this zoom is very different from any other zoom that is non-symmetric with respect to $c \mapsto \overline{c}$. More importantly, with a keen eye for geometry, one can deduce exactly where in the boundary of $\mathcal{M}$ this zoom is, and, more importantly, what the corresponding dynamical behavior in the associated bulb is. It turns out that there are several different geometric and dynamical ways to understand the structure of these bulbs. We will look at this geometrically, but the real way to understand this uses techniques from complex analysis.

For simplicity, let's concentrate on the bulbs attached to the main cardioid. How do we know what their period is? One way is easy: look at the bulb. There is an antenna attached to this bulb. This antenna has a junction point from which a certain number of spokes emanate. The number of these spokes tells us exactly what the

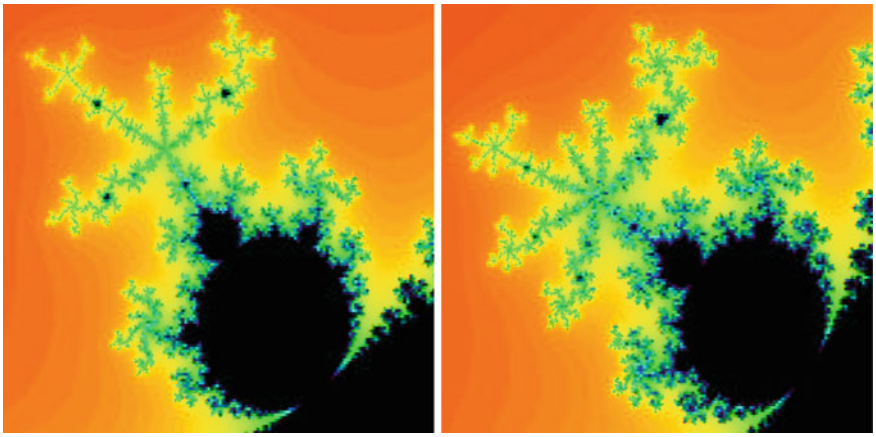


Fig. 3 Period 5 and 7 bulbs hanging off the main cardioid

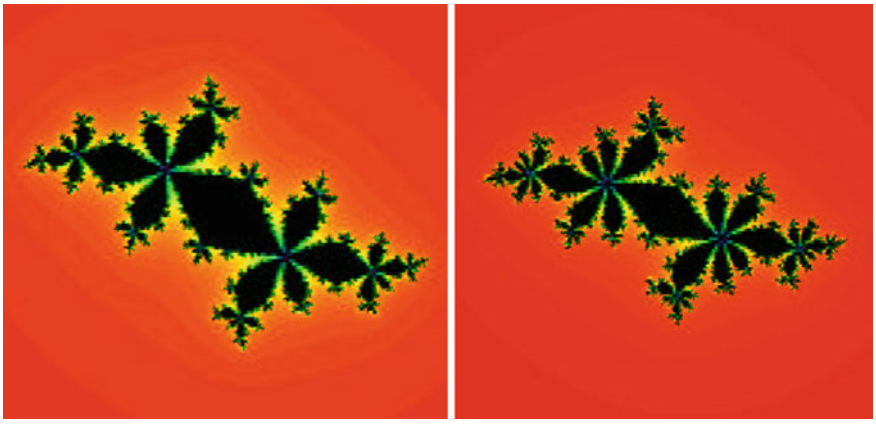


Fig. 4 Julia sets drawn from the above period 5 and 7 bulbs hanging off the main cardioid. Note that there are 4 and 6 "ears" hanging off the central disks of these filled Julia sets

period is. For example, in Fig. 3, we display two bulbs having periods 5 and 7. Note that this is the exact number of antennas hanging off the junction point in the antenna of each bulb.

There is another way to read off the periods of these bulbs. Choose a parameter from the interior of a period n bulb and plot the corresponding filled Julia set. There is a central disk in these filled Julia sets that surrounds the origin. Then there are exactly $n - 1$ smaller disks that join this main disk at certain junction points. For example, in Fig. 2, we see that the rabbit has two "ears" attached to the central disk and the period of this bulb is $2 + 1 = 3$. Similarly, the basilica has just 1 ear and the period here is $1 + 1 = 2$. In Fig. 4, we display Julia sets from the above period 5 and period 7 bulbs, and we see the same phenomenon.

Now let us turn to the arrangement of the bulbs around the main cardioid. To do this, we assign a fraction p/q to each of these bulbs. Here q is the period of the bulb, so the question is: what is p? There are several geometric and dynamical ways to determine p. Look at the period five bulb in Fig. 3. We call the spoke of the antenna that extends down to the bulb from the junction point the *principal spoke*. Note that the "shortest" spoke (that is not the principal spoke) is located $2/5$ of a turn in the counter-clockwise direction from the principal spoke. And this bulb is then the $2/5$-bulb. In that same figure, we also see that the period 7-bulb is, in fact, the $3/7$-bulb.

A second way to see this is to turn to the filled Julia set. In Fig. 4, each of the filled Julia sets has a main component that surrounds the origin together with $q-1$ ears attached at one point. Note where the "smallest" ear is located; it is exactly p/q of a turn in the counterclockwise direction from main component.

And then there is a third way to read off p/q. Simply plot the points on the attracting cycle of period q in the Fatou set. What you see is that this cycle moves around the ears and the main component, rotating by p/q of a turn at each stage. So there is a very nice connection between the geometry of the Mandelbrot set and Julia sets and the dynamics of P_c.

One curious fact that relates to the Farey tree involves the size of the bulbs hanging off the main cardioid. To begin, we think of the root point of the main cardioid as being the cusp at $c = 1/4$. Then we call the main cardioid the $0/1$-bulb. The root point of any other bulb is just the point where this bulb is attached to the main cardioid. Now which is the largest bulb between the root points of the $0/1$ and $1/2$-bulbs (in, say, the upper portion of $\mathcal{M}$)? It is clearly the $1/3$-bulb. And note that $1/3$ is obtained from the previous two fractions by *Farey addition*, i.e., adding the numerators and adding the denominators

$$\frac{0}{1}\text{“}+\text{”}\frac{1}{2}=\frac{1}{3}.$$

Similarly, the largest bulb between the $1/3$ and $1/2$-bulbs is the $2/5$-bulb, again given by Farey addition.

$$\frac{1}{3}\text{“}+\text{”}\frac{1}{2}=\frac{2}{5}.$$

And the largest bulb between the $2/5$ and $1/2$-bulb is the $3/7$-bulb while the largest bulb between the $2/5$ and $1/3$-bulbs is the $3/8$-bulb and so on along the "Farey tree".

Then it follows that these bulbs are arranged around the boundary of the main cardioid in the exact order of the rational numbers in the unit interval. Actually, techniques from calculus can be used to prove this fairly easily. For more details, see [1–3]. An online, interactive discussion of this (with plenty of animations) called the Mandelbrot Set Explorer is available at http://math.bu.edu/DYSYS/explorer.

Using similar techniques from geometry, one can identify the other sub-bulbs in the Mandelbrot set. Unfortunately, there are many other points in the Mandelbrot set that this approach does not apply to; indeed, despite the simplicity of the function z^2+c, there are still many c-values in $\mathcal{M}$ for which we have no idea what is happening

in the corresponding Julia set and what is the nearby structure in the Mandelbrot set. For example, along the boundary of the main cardioid we have only looked at the parameters corresponding to "rational" root points as discussed above. But there are uncountably many other points along the boundary of the cardioid. These correspond to "irrational" points. We understand the behavior of P_c at the so-called "highly" irrational points, but the parameters at the "not-so-irrational" points have behavior that is still not understood. This is one of the major open problems in this area of mathematics. For a basic introduction to complex dynamics, see [4]. A more advanced survey of this field is John Milnor's book [5].

References

1. Devaney, R.L., Moreno Rocha, M.: Geometry of the antennas in the Mandelbrot set. Fractals **10**, 39–46 (2002)
2. Devaney, R.L., Moreno Rocha, M.: The fractal geometry of the Mandelbrot set: I. periods of the bulbs. In: Fractals, Graphics, and Mathematics Education. MAA Notes, vol. 58, pp. 61–68 (2002)
3. Devaney, R.L., Moreno Rocha, M.: The fractal geometry of the Mandelbrot set: II. How to add and how to count. Fractals **3**, 629–640 (1995)
4. Devaney, R.L.: An Introduction to Chaotic Dynamical Systems, 2nd edn. Westview Press, Boulder (2003)
5. Milnor, J.: Dynamics in One Complex Variable. Princeton University Press, Princeton (2006)

Is There a World Behind Shannon? Entropies for Complex Systems

Stefan Thurner and Rudolf Hanel

Abstract In their seminal works, Shannon and Khinchin showed that assuming four information theoretic axioms the entropy must be of Boltzmann-Gibbs type, $S = -\sum_i p_i \log p_i$. In many physical systems one of these axioms may be violated. For non-ergodic systems the so called separation axiom (Shannon-Khinchin axiom 4) is not valid. We show that whenever this axiom is violated the entropy takes a more general form, $S_{c,d} \propto \sum_i^W \Gamma(d+1, 1-c\log p_i)$, where c and d are scaling exponents and $\Gamma(a,b)$ is the incomplete gamma function. These exponents (c,d) define equivalence classes for *all*!, interacting and non interacting, systems and unambiguously characterize any statistical system in its thermodynamic limit. The proof is possible because of two newly discovered scaling laws which any entropic form has to fulfill, if the first three Shannon-Khinchin axioms hold [1]. (c,d) can be used to define equivalence classes of statistical systems. A series of known entropies can be classified in terms of these equivalence classes. We show that the corresponding distribution functions are special forms of Lambert-$\mathcal{W}$ exponentials containing—as special cases—Boltzmann, stretched exponential, and Tsallis distributions (power-laws). We go on by showing how the dependence of phase space volume $W(N)$ of a classical system on its size N, uniquely determines its extensive entropy, and in particular that the requirement of extensivity fixes the exponents (c,d), [2]. We

PACS 05.20.-y—Classical statistical mechanics
PACS 02.50.Cw—Probability theory
PACS 05.90.+m—Other topics in statistical physics, thermodynamics, and nonlin. dyn. systems.

S. Thurner (✉) · R. Hanel
Section for Science of Complex Systems, Medical University of Vienna,
Spitalgasse 23, A-1090 Vienna, Austria
e-mail: stefan.thurner@meduniwien.ac.at

S. Thurner
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM87501, USA

R. Hanel
e-mail: rudolf.hanel@meduniwien.ac.at

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_2,

give a concise criterion when this entropy is not of Boltzmann-Gibbs type but has to assume a *generalized* (non-additive) form. We showed that generalized entropies can only exist when the dynamically (statistically) relevant fraction of degrees of freedom in the system vanishes in the thermodynamic limit [2]. These are systems where the bulk of the degrees of freedom is frozen and is practically statistically inactive. Systems governed by generalized entropies are therefore systems whose phase space volume effectively collapses to a lower-dimensional 'surface'. We explicitly illustrated the situation for binomial processes and argue that generalized entropies could be relevant for self organized critical systems such as sand piles, for spin systems which form meta-structures such as vortices, domains, instantons, etc., and for problems associated with anomalous diffusion [2]. In this contribution we largely follow the lines of thought presented in [1–3].

Keywords Scaling of entropy · (c, d) entropy · Axiomatic derivation of entropy · Extensivity · Non-ergodic systems · Entropy for non-Markovian processes

1 Introduction

In their seminal works, Shannon and Khinchin showed that assuming four information theoretic axioms the entropy must be of Boltzmann-Gibbs type, $S = -\sum_i p_i \log p_i$. In many physical systems one of these axioms may be violated. For non-ergodic systems the so called separation axiom (Shannon-Khinchin axiom 4) is not valid. We show that whenever this axiom is violated the entropy takes a more general form, $S_{c,d} \propto \sum_i^W \Gamma(d+1, 1-c\log p_i)$, where c and d are scaling exponents and $\Gamma(a, b)$ is the incomplete gamma function. These exponents (c, d) define equivalence classes for *all*!, interacting and non interacting, systems and unambiguously characterize any statistical system in its thermodynamic limit. The proof is possible because of two newly discovered scaling laws which any entropic form has to fulfill, if the first three Shannon-Khinchin axioms hold [1]. (c, d) can be used to define equivalence classes of statistical systems. A series of known entropies can be classified in terms of these equivalence classes. We show that the corresponding distribution functions are special forms of Lambert-$\mathcal{W}$ exponentials containing—as special cases—Boltzmann, stretched exponential, and Tsallis distributions (power-laws). We go on by showing how the dependence of phase space volume $W(N)$ of a classical system on its size N, uniquely determines its extensive entropy, and in particular that the requirement of extensivity fixes the exponents (c, d), [2]. We give a concise criterion when this entropy is not of Boltzmann-Gibbs type but has to assume a *generalized* (non-additive) form. We showed that generalized entropies can only exist when the dynamically (statistically) relevant fraction of degrees of freedom in the system vanishes in the thermodynamic limit [2]. These are systems where the bulk of the degrees of freedom is frozen and is practically statistically inactive. Systems governed by generalized entropies are therefore systems whose phase space volume effectively collapses to a lower-dimensional 'surface'. We explicitly

illustrated the situation for binomial processes and argue that generalized entropies could be relevant for self organized critical systems such as sand piles, for spin systems which form meta-structures such as vortices, domains, instantons, etc., and for problems associated with anomalous diffusion [2]. In this contribution we largely follow the lines of thought presented in [1–3].

Theorem number 2 in the seminal 1948 paper, *The Mathematical Theory of Communication* [4], by Claude Shannon, proves the existence of the one and only form of entropy, given that three fundamental requirements hold. A few years later A.I. Khinchin remarked in his *Mathematical Foundations of Information Theory* [5]: "However, Shannon's treatment is not always sufficiently complete and mathematically correct so that, besides having to free the theory from practical details, in many instances I have amplified and changed both the statement of definitions and the statement of proofs of theorems." Khinchin adds a fourth axiom. The three fundamental requirements of Shannon, in the 'amplified' version of Khinchin, are known as the Shannon-Khinchin (SK) axioms. These axioms list the requirements needed for an entropy to be a reasonable measure of the 'uncertainty' about a finite probabilistic system. Khinchin further suggests to also use entropy as a measure of the information *gained* about a system when making an 'experiment', i.e. by observing a realization of the probabilistic system.

Khinchin's first axiom states that for a system with W potential outcomes (states) each of which is given by a probability $p_i \geq 0$, with $\sum_{i=1}^{W} p_i = 1$, the entropy $S(p_1, \ldots, p_W)$ as a measure of uncertainty about the system must take its maximum for the equi-distribution $p_i = 1/W$, for all i.

Khinchin's second axiom (missing in [4]) states that any entropy should remain invariant under adding zero-probability states to the system, i.e. $S(p_1, \ldots, p_W) = S(p_1, \ldots, p_W, 0)$.

Khinchin's third axiom (separability axiom) finally makes a statement of the composition of two finite probabilistic systems A and B. If the systems are independent of each other, entropy should be additive, meaning that the entropy of the combined system $A+B$ should be the sum of the individual systems, $S(A + B) = S(A)+S(B)$. If the two systems are dependent on each other, the entropy of the combined system, i.e. the information given by the realization of the two finite schemes A and B, $S(A + B)$, is equal to the information gained by a realization of system A, $S(A)$, plus the mathematical expectation of information gained by a realization of system B, after the realization of system A, $S(A + B) = S(A) + S|_A(B)$.

Khinchin's fourth axiom is the requirement that entropy is a continuous function of all its arguments p_i and does not depend on anything else.

Given these axioms, the *Uniqueness theorem* [5] states that the one and only possible entropy is

$$S(p_1, \ldots, p_W) = -k \sum_{i=1}^{W} p_i \log p_i, \tag{1}$$

where k is an arbitrary positive constant. The result is of course the same as Shannon's. We call the combination of 4 axioms the Shannon-Khinchin (SK) axioms.

From information theory now to physics, where systems may exist that violate the separability axiom. This might especially be the case for non-ergodic, complex systems exhibiting long-range and strong interactions. Such complex systems may show extremely rich behavior in contrast to simple ones, such as gases. There exists some hope that it should be possible to understand such systems also on a thermodynamical basis, meaning that a few measurable quantities would be sufficient to understand their macroscopic phenomena. If this would be possible, through an equivalent to the second law of thermodynamics, some appropriate entropy would enter as a fundamental concept relating the number of microstates in the system to its macroscopic properties. Guided by this hope, a series of so called generalized entropies have been suggested over the past decades, see [6–11] and Table 1. These entropies have been designed for different purposes and have not been related to a fundamental origin. Here we ask how generalized entropies can look like if they fulfill some of the Shannon-Khinchin axioms, but explicitly violate the separability axiom. We do this axiomatically as first presented in [1]. By doing so we can relate a large class of generalized entropies to a single fundamental origin.

The reason why this axiom is violated in some physical, biological or social systems is *broken ergodicity*, i.e. that not all regions in phase space are visited and many micro states are effectively 'forbidden'. Entropy relates the number of micro states of a system to an *extensive* quantity, which plays the fundamental role in the systems thermodynamical description. Extensive means that if two initially isolated, i.e. sufficiently separated systems, A and B, with W_A and W_B the respective numbers of states, are brought together, the entropy of the combined system $A + B$ is $S(W_{A+B}) = S(W_A) + S(W_B)$. W_{A+B} is the number of states in the combined system $A + B$. This is not to be confused with *additivity* which is the property that $S(W_A W_B) = S(W_A) + S(W_B)$. Both, extensivity and additivity coincide if number of states in the combined system is $W_{A+B} = W_A W_B$. Clearly, for a non-interacting system Boltzmann-Gibbs-Shannon entropy, $S_{\mathrm{BG}}[p] = -\sum_i^W p_i \ln p_i$, is extensive *and* additive. By 'non-interacting' (short-range, ergodic, sufficiently mixing, Markovian, ...) systems we mean $W_{A+B} = W_A W_B$. For interacting statistical systems the latter is in general not true; phase space is only partly visited and $W_{A+B} < W_A W_B$. In this case, an additive entropy such as Boltzmann-Gibbs-Shannon can no longer be extensive and vice versa. To ensure extensivity of entropy, an entropic form should be found for the particular interacting statistical systems at hand. These entropic forms are called *generalized entropies* and usually assume trace form [6–11]

$$S_g[p] = \sum_{i=1}^{W} g(p_i), \tag{2}$$

W being the number of states. Obviously not all generalized entropic forms are of this type. Rényi entropy e.g. is of the form $G(\sum_i^W g(p_i))$, with G a monotonic function. We use trace forms Eq. (2) for simplicity. Rényi forms can be studied in exactly the same way as will be shown, however at more technical cost.

Table 1 Order in the zoo of recently introduced entropies for which SK1-SK3 hold. All of them are special cases of the entropy given in Eq. (3) and their asymptotic behavior is uniquely determined by c and d. It can be seen immediately that $S_{q>1}$, S_b and S_E are asymptotically identical; so are $S_{q<1}$ and S_κ, as well as S_η and S_γ

Entropy		c	d	Reference
$S_{c,d} = er\sum_i \Gamma(d+1, 1-c\ln p_i) - cr$	$(r = (1-c+cd)^{-1})$	c	d	
$S_{BG} = \sum_i p_i \ln(1/p_i)$		1	1	[5]
$S_{q<1}(p) = \frac{1-\sum p_i^q}{q-1}$	$(q < 1)$	$c = q < 1$	0	[6]
$S_\kappa(p) = -\sum_i p_i \frac{p_i^\kappa - p_i^{-\kappa}}{2\kappa}$	$(0 < \kappa \leq 1)$	$c = 1-\kappa$	0	[8]
$S_{q>1}(p) = \frac{1-\sum p_i^q}{q-1}$	$(q > 1)$	1	0	[6]
$S_b(p) = \sum_i (1-e^{-bp_i}) + e^{-b} - 1$	$(b > 0)$	1	0	[9]
$S_E(p) = \sum_i p_i (1-e^{\frac{p_i-1}{p_i}})$		1	0	[10]
$S_\eta(p) = \sum_i \Gamma(\frac{\eta+1}{\eta}, -\ln p_i) - p_i\Gamma(\frac{\eta+1}{\eta})$	$(\eta > 0)$	1	$d = \frac{1}{\eta}$	[7]
$S_\gamma(p) = \sum_i p_i \ln^{1/\gamma}(1/p_i)$		1	$d = 1/\gamma$	[14], footnote 11, page 60
$S_\beta(p) = \sum_i p_i^\beta \ln(1/p_i)$		$c = \beta$	1	[15]

Let us revisit the Shannon-Khinchin axioms in the light of generalized entropies of trace form Eq. (2). Specifically axioms SK1-SK3 (now re-ordered) have implications on the functional form of g

- SK1: The requirement that S depends continuously on p implies that g is a continuous function.
- SK2: The requirement that the entropy is maximal for the equi-distribution $p_i = 1/W$ (for all i) implies that g is a concave function.
- SK3: The requirement that adding a zero-probability state to a system, $W+1$ with $p_{W+1} = 0$, does not change the entropy, implies that $g(0) = 0$.
- SK4 (separability axiom): The entropy of a system—composed of sub-systems A and B—equals the entropy of A plus the expectation value of the entropy of B, conditional on A. Note that this also corresponds exactly to Markovian processes.

As mentioned, if SK1 to SK4 hold, the only possible entropy is the Boltzmann-Gibbs-Shannon entropy. We are now going to derive the extensive entropy when the separability axiom SK4 is violated. Obviously this entropy will be more general and should contain BG entropy as a special case.

We now assume that axioms SK1, SK2, SK3 hold, i.e. we restrict ourselves to trace form entropies with g continuous, concave and $g(0) = 0$. These systems we call *admissible* systems. Admissible systems when combined with a maximum entropy principle show remarkably simple mathematical properties [12, 13].

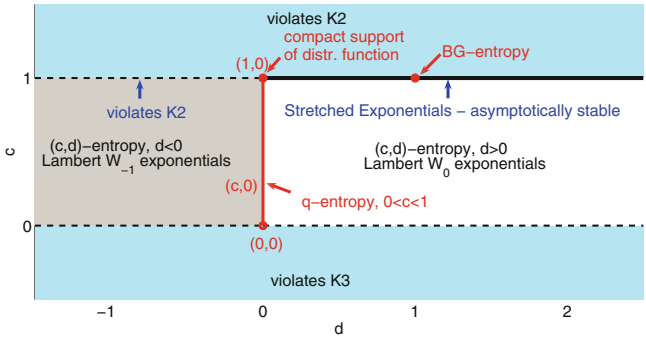


Fig. 1 Entropies parametrized in the (c, d)-plane, with their associated distribution functions. BG entropy corresponds to $(1, 1)$, Tsallis entropy to $(c, 0)$, and entropies for stretched exponentials to $(1, d > 0)$. Entropies leading to distribution functions with compact support, belong to equivalence class $(1, 0)$. Figure from [3]

This generalized entropy for (large) admissible statistical systems (SK1-SK3 hold) is derived from two hitherto unexplored fundamental scaling laws of extensive entropies [1]. Both scaling laws are characterized by exponents c and d, respectively, which allow to uniquely define equivalence classes of entropies, meaning that two entropies are equivalent in the thermodynamic limit if their exponents (c, d) coincide. Each admissible system belongs to one of these equivalence classes (c, d), [1].

In terms of the exponents (c, d) we showed in [1] that all generalized entropies have the form

$$S_{c,d} \propto \sum_i^W \Gamma(d+1, 1 - c \log p_i) \tag{3}$$

with $\Gamma(a, b) = \int_b^\infty dt\, t^{a-1} \exp(-t)$ the incomplete Gamma-function.

1.1 Special Cases of Entropic Equivalence Classes

Let us look at some specific equivalence classes (c, d)

- Boltzmann-Gibbs entropy belongs to the $(c, d) = (1, 1)$ class. One gets from Eq. (3)

$$S_{1,1}[p] = \sum_i g_{1,1}(p_i) = -\sum_i p_i \ln p_i + 1 \quad . \tag{4}$$

- Tsallis entropy belongs to the $(c, d) = (c, 0)$ class. From Eq. (3) and the choice $r = 1/(1 - c)$ (see below) we get

$$S_{c,0}[p] = \sum_i g_{c,0}(p_i) = \frac{1-\sum_i p_i^c}{c-1} + 1\,. \tag{5}$$

Note, that although the *pointwise* limit $c \to 1$ of Tsallis entropy yields BG entropy, the asymptotic properties $(c, 0)$ do *not* change continuously to $(1, 1)$ in this limit! In other words the thermodynamic limit and the limit $c \to 1$ do not commute.

- The entropy related to stretched exponentials [7] belongs to the $(c, d) = (1, d)$ classes, see Table 1. As a specific example we compute the $(c, d) = (1, 2)$ case,

$$S_{1,2}[p] = 2\left(1 - \sum_i p_i \ln p_i\right) + \frac{1}{2}\sum_i p_i \,(\ln p_i)^2\,, \tag{6}$$

leading to a superposition of two entropy terms, the asymptotic behavior being dominated by the second.

Other entropies which are special cases of our scheme are found in Table 1.

Inversely, for any given entropy we are now in the remarkable position to characterize *all* large SK1-SK3 systems by a pair of two exponents (c, d), see Fig. 1. For example, for $g_{\rm BG}(x) = -x\ln(x)$ we have $c = 1$, and $d = 1$. $S_{\rm BG}$ therefore belongs to the universality class $(c, d) = (1, 1)$. For $g_q(x) = (x - x^q)/(1 - q)$ (Tsallis entropy) and $0 < q < 1$ one finds $c = q$ and $d = 0$, and Tsallis entropy, S_q, belongs to the universality class $(c, d) = (q, 0)$. Other examples are listed in Table 1.

The universality classes (c, d) are equivalence classes with the equivalence relation given by: $g_\alpha \equiv g_\beta \Leftrightarrow c_\alpha = c_\beta$ and $d_\alpha = d_\beta$. This relation partitions the space of all admissible g into equivalence classes completely specified by the pair (c, d).

2 Distribution Functions

Distribution functions associated with our Γ-entropy, Eq. (3), can be derived from so-called generalized logarithms of the entropy. Under the maximum entropy principle (given ordinary constraints) the inverse functions of these logarithms, $\mathcal{E} = \Lambda^{-1}$, are the distribution functions, $p(\epsilon) = \mathcal{E}_{c,d,r}(-\epsilon)$, where for example r can be chosen $r = (1 - c + cd)^{-1}$. One finds [1]

$$\mathcal{E}_{c,d,r}(x) = e^{-\frac{d}{1-c}\left[\mathcal{W}_k\left(B(1-x/r)^{\frac{1}{d}}\right) - \mathcal{W}_k(B)\right]}\,, \tag{7}$$

with the constant $B \equiv \frac{(1-c)r}{1-(1-c)r}\exp\left(\frac{(1-c)r}{1-(1-c)r}\right)$. The function $\mathcal{W}_k$ is the k'th branch of the Lambert-$\mathcal{W}$ function which—as a solution to the equation $x = \mathcal{W}(x)\exp(\mathcal{W}(x))$—has only two real solutions W_k, the branch $k = 0$ and branch $k = -1$. Branch $k = 0$ covers the classes for $d \geq 0$, branch $k = -1$ those for $d < 0$.

2.1 Special Cases of Distribution Functions

It is easy to verify that the class $(c, d) = (1, 1)$ leads to Boltzmann distributions, and the class $(c, d) = (c, 0)$ yields power-laws, or more precisely, Tsallis distributions i.e. q-exponentials. All classes associated with $(c, d) = (1, d)$, for $d > 0$ are associated with stretched exponential distributions. Expanding the $k = 0$ branch of the Lambert-$\mathcal{W}$ function $W_0(x) \sim x - x^2 + \cdots$ for $1 \gg |x|$, the limit $c \to 1$ is shown to be a stretched exponential. It was shown that r does not effect its asymptotic properties (tail of the distributions), but can be used to incorporate finite size properties of the distribution function for small x.

3 How to Determine the Exponents c and d?

In [2] we have shown that the requirement of extensivity determines uniquely both exponents c and d. What does extensivity mean? Consider a system with N elements. The number of system configurations (microstates) as a function of N are denoted by $W(N)$. Starting with SK2, $p_i = 1/W$ (for all i), we have $S_g = \sum_{i=1}^{W} g(p_i) = W g(1/W)$. As mentioned above extensivity for two subsystems A and B means that

$$W_{A+B} g\,(1/W_{A+B}) = W_A g\,(1/W_A) + W_B g\,(1/W_B). \tag{8}$$

Using this equation one can straight forwardly derive the formulas (for details see [2])

$$\frac{1}{1-c} = \lim_{N\to\infty} N \frac{W'(N)}{W(N)}. \tag{9}$$

$$d = \lim_{N\to\infty} \log W \left(\frac{1}{N} \frac{W}{W'} + c - 1 \right). \tag{10}$$

Here W' means the derivative with respect to N.

3.1 A Note on Rényi-Type Entropies

Rényi entropy is obtained by relaxing SK4 to the unconditional additivity condition. Following the same scaling idea for Rényi-type entropies, $S = G(\sum_{i=1}^{W} g(p_i))$, with G and g some functions, one gets

$$\lim_{W\to\infty} \frac{S(\lambda W)}{S(W)} = \lim_{s\to\infty} \frac{G\left(\lambda f_g(\lambda^{-1}) s\right)}{G(s)}, \tag{11}$$

where $f_g(z) = \lim_{x\to 0} g(zx)/g(x)$. The expression $f_G(s) \equiv \lim_s G(sy)/G(s)$, provides the starting point for deeper analysis which now gets more involved. In particular, for Rényi entropy with $G(x) \equiv \ln(x)/(1-\alpha)$ and $g(x) \equiv x^\alpha$, the asymptotic properties yield the class $(c, d) = (1, 1)$, (BG entropy) meaning that Rényi entropy is additive. However, in contrast to the trace form entropies used above, Rényi entropy can be shown to be *not* Lesche stable, as was observed before [16–20]. All of the $S = \sum_i^W g(p_i)$ entropies can be shown to be Lesche stable, see [3].

4 Discussion

We discuss recently discovered scaling laws for trace form entropies for systems that fulfill the first three Shannon-Khinchin axioms. In analogy to critical exponents these laws are characterized by two scaling exponents (c, d), which define generalized entropies. We showed that a particular entropic form—parametrized by these two exponents—covers all *admissible* systems (Shannon-Khinchin axioms 1–3 hold, 4 is violated). In other words every statistical system has its pair of unique exponents in the large size limit, its entropy is then given by $S_{c,d} \sim \sum_i^W \Gamma\,(1+d\,,\; 1-c\ln p_i)$. The requirement of extensivity uniquely determines the scaling exponents c and d in terms of the growth of phase space as a function of system size.

The exponents for BG systems are $(c, d) = (1, 1)$, systems characterized by stretched exponentials belong to the class $(c, d) = (1, d)$, and Tsallis systems have $(c, d) = (q, 0)$. In the context of a maximum entropy principle, the associated distribution functions of *all* systems (c, d) are shown to belong to a class of exponentials involving Lambert-$\mathcal{W}$ functions, given in Eq. (7). There are no other options for tails in distribution functions other than these.

The equivalence classes characterized by the exponents (c, d) form *basins of asymptotic equivalence*. In general these basins characterize interacting statistical (non-additive) systems. There exists an analogy between these basins of asymptotic *equivalence* and the *basin of attraction* of weakly interacting, uncorrelated systems subject to the law of large numbers, i.e. the central limit theorem. Any system within a given equivalence class may show individual characteristics as long as it is small. Systems belonging to the same class will start behaving similarly as they become larger, and in the thermodynamic limit they become identical. Distribution functions converge to those uniquely determined by (c, d). A further interesting feature of all admissible systems is that they all are *Lesche stable*. The proof is found in [3].

Finally, the classification scheme for generalized entropies of type $S = \sum_i g(p_i)$ can be extended to entropies of e.g. Rényi type, i.e. $S = G(\sum_i g(p_i))$, see [3].

References

1. Hanel, R., Thurner, S.: Europhys. Lett. **93**, 20006 (2011)
2. Hanel, R., Thurner, S.: Europhys. Lett. **96**, 50003 (2011)
3. Thurner, S., Hanel, R.: Recent Advances in Generaliized Information Measures and Statistics. Bentham Science eBook, (in production 2013)
4. Shannon, C.E.: Bell Syst. Tech. J. **27**:379,623 (1948)
5. Khinchin, A.I.: Mathematical Foundations of Information Theory. Dover Publications, New York (1957)
6. Tsallis, C.: J. Stat. Phys. **52**, 479 (1988)
7. Anteneodo, C., Plastino, A.R.: J. Phys. A: Math. Gen. **32**, 1089 (1999)
8. Kaniadakis, G.: Phys. Rev. E **66**, 056125 (2002)
9. Curado, E.M.F., Nobre, F.D.: Phys. A **335**, 94 (2004)
10. Tsekouras, G.A., Tsallis, C.: Phys. Rev. E **71**, 046144 (2005)
11. Hanel, R., Thurner, S.: Phys. A **380**, 109 (2007)
12. Hanel, R., Thurner, S., Gell-Mann, M.: PNAS **108**, 6390–6394 (2011)
13. Hanel, R., Thurner, S., Gell-Mann, M.: PNAS **109**, 19151–19154 (2012)
14. Tsallis, C.: Introduction to Nonextensive Statistical Mechanics. Springer, New York (2009)
15. Shafee, F.: IMA J. Appl. Math. **72**, 785 (2007)
16. Lesche, B.: J. Stat. Phys. **27**, 419 (1982)
17. Abe, S.: Phys. Rev. E **66**, 046134 (2002)
18. Jizba, P., Arimitsu, T.: Phys. Rev. E **69**, 026128 (2004)
19. Kaniadakis, G., Scarfone, A.M.: Phys. A **340**, 102 (2004)
20. Hanel, R., Thurner, S., Tsallis, C.: Europhys. Lett. **85**, 20005 (2009)

Complex Systems Science: From Cell Regulation to the Global Food Crisis

Yaneer Bar-Yam

Abstract Insights and methods of complex systems science are transforming science and providing clarity about the impact of policies to address major societal problems. These conceptual and mathematical advances allow us to study interdependence, patterns, networks, multiscale behaviors, and "big data." Here I focus on the application of these advances to real-world concerns. I discuss case studies from global socioeconomic systems and immune cell regulation. Our analysis of the global food crisis exposes the causes and consequences of rapidly increasing and volatile food prices. Food price spikes in 2007–2008 and 2010–2011 triggered food riots across the world and precipitated the Arab Spring. Our quantitative models of nonequilibrium markets show that the food price increases are due to (1) US biofuel quotas increasing the amount of corn to ethanol conversion and (2) deregulation of commodity trading enabling speculator trend-following to cause bubbles and crashes. Policy action by the US and the European Union could alleviate or even resolve these problems. Our analysis of cell regulation makes use of gene expression data to obtain whole-cell regulatory models describing the response of immune cells to dynamic perturbations. Moreover, we have shown that cell dynamics are controlled by attractor states with implications for understanding biological development and treating cancer. Our analyses demonstrate the opportunity for complex systems science to inform both social policy decisions and medical advances.

Keywords Complex systems, Economic crisis, Cell regulation

Y. Bar-Yam (✉)
New England Complex Systems Institute, Cambridge, MA, USA
e-mail: yaneer@necsi.edu

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_3,

1 Vulnerability of Financial Systems

When an economic system is robust it can function under a variety of stresses, and when it is not robust even minor perturbations cause cascading failures and dislocation of its essential function. We have seen evidence of such dislocation in the failures of the US financial system since 2007 and in government financial and economic rescue actions that only provide temporary and insufficient results.

Our results show that government policy decisions, often in deregulation but also in regulation, have undermined the ability of our economic system to function and made it highly susceptible to crises. The economic system is an essential part of the functioning of our society. When functioning properly its action enables both basic survival and many other opportunities for us, individually and collectively. The need for its functional reliability should be apparent, and should not be assumed. Similar to other systems with emergent behaviors, economic activity depends on a framework in which it can function successfully. Our analysis is designed to identify key aspects of the framework that enable economic activity, and without which systemic failure is likely. In each case we identify the nature of the systemic dysfunction associated with regulatory changes.

Our research has identified three areas of regulatory activity that have played a major role in the global financial, economic and food crises.

1.1 Banking Deregulation

The repeal of the Glass-Steagall Act of 1934 by the Gramm-Leach-Bliley Act in 1999 allowed the creation of the "too big to fail" banks. A lobbying effort by Citigroup to achieve this legislation is a matter of public record. The increased vulnerability of the economy due to this deregulation has been widely discussed and is analyzed in our paper "Networks of Economic Market Interdependence and Systemic Risk" [1]. In particular, the banks facilitated the propagation of the mortgage crisis to the economy as a whole (Fig. 1). We conclude that reinstating banking regulation that separates domains of banking activity would inhibit cascading economic failures and play an important role in restoring economic stability.

1.2 Commodity Futures Deregulation

Our research shows that rapid increases in food and other commodity prices during this past decade have their basis in deregulation of the commodities markets. This is described in our paper "The Food Crises: A Quantitative Model of Food Prices Including Speculators and Ethanol Conversion" [2]. The repeal of the Commodity Exchange Act of 1936 by the Commodity Futures Modernization Act of 2000 removed limits on trading. These trading limits were established in order to limit the impact of speculation. Their repeal enabled rapid growth in commodity

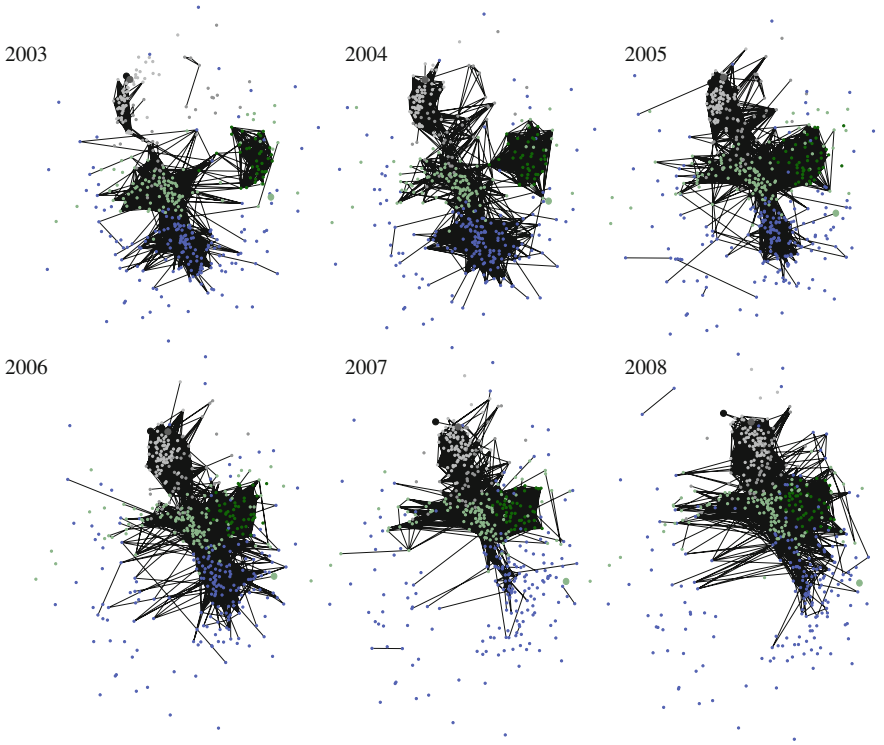


Fig. 1 Network of correlations of market daily returns for years indicated. Dots represent individual corporations colored according to economic sector: technology (*blue*), basic materials including oil companies (*light grey*) and others (*dark grey*), and finance including real-estate (*dark green*) and other (*light green*). Links are shown for pairs of stocks with correlation above a certain level. The network exhibits increasing integration across sectors over time. The banking sector links parts of the economy, transmitting the collapse of the real-estate market to the rest

index funds and contributed to the volume of speculative activity that resulted in bubbles and crashes that disrupt supply and demand economics in these systems (see Fig. 2). When supply and demand are not in equilibrium, the effectiveness and even stability of the economic system are undermined. The increases in food prices have resulted in a global food crisis, causing widespread suffering and major food riots. The Commodity Futures Trading Commission is currently in process of reinstating position limits, and the primary corporation meeting with them is Goldman Sachs, which also manages the largest commodity index fund.

1.3 Stock Market Deregulation

The repeal by the SEC in July 2007 of the "uptick rule" of 1938 eliminated its intended role in stabilizing the stock markets against the rapid selling of borrowed shares, whether for illegal intentional manipulation or otherwise. Our research provided

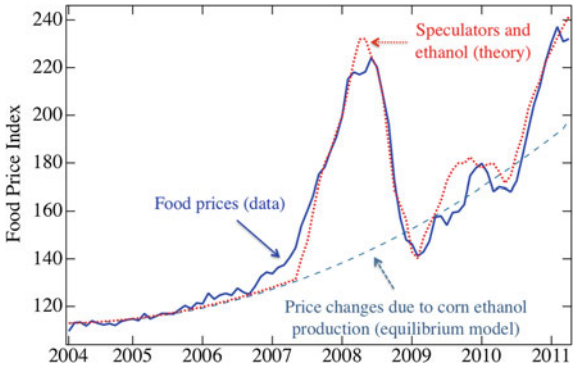


Fig. 2 Actual food prices and quantitative dynamic model of commodity speculation and mandated use of corn in ethanol production. The corn to ethanol conversion causes price increases due to its large impact on supply and demand, and speculation causes bubbles in 2007–2008, and 2010–2011

a scientific analysis of the impact of the uptick rule [3] (see also Fig. 3). Despite overwhelming public and professional support for reinstatement of the uptick rule, the SEC listened to industry advocates, especially hedge funds, who claimed that the uptick rule was not beneficial. In March of 2011 the SEC implemented a rule that is only in effect if a stock's price decreases by 10 % in a single day. Our analysis shows that this rule does not have the same stabilizing effect as the original uptick rule. The "flash crash" on May 6, 2010 and mini flash crashes manifest market dysfunction, and they are consistent with our analysis of the consequences of this deregulation.

2 Complexity in Biological Systems

2.1 Attractors and Democratic Dynamics

The functional identity of a cell is largely determined by the regulated expression (transcription) of thousands of genes, so how it maintains a particular transcriptional state is of critical importance. Developmental biologists study how embryonic cells navigate a series of intermediate transcriptional states before settling into a final adult state; microbiologists identify the mechanisms by which transcription is altered by environmental perturbation; and oncologists seek to identify how cells switch from benign to cancerous. Consider two concepts of transcriptional regulation. In a "molecular autocracy" master genes respond to environmental or developmental stimuli by regulating thousands of genes, either directly or through other transcription factors. In a "molecular democracy" all genes exert a regulatory influence on all other genes, and phenotypic change (altered cell behavior) is brought about through the concerted action of thousands of genes. These scenarios are extreme and cells operate under a condition that is somewhere intermediate (see Fig. 4). But the choice of concept affects how regulation is studied.

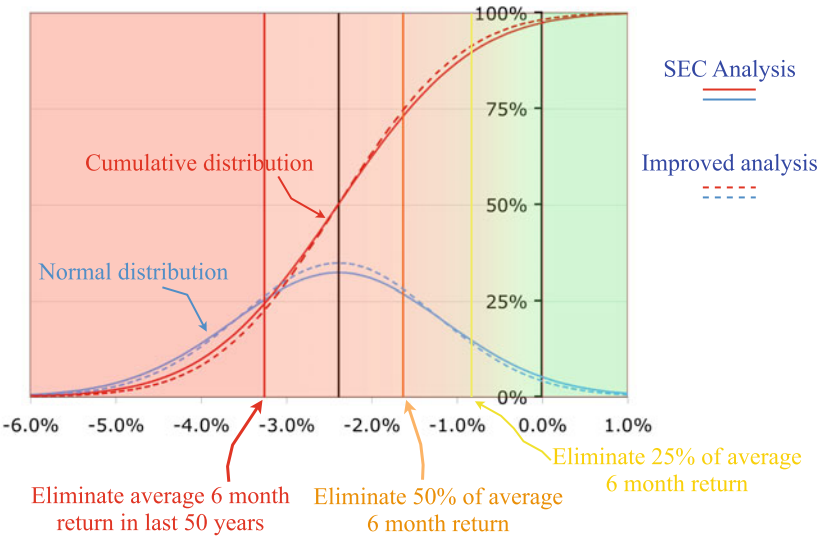


Fig. 3 Probability distribution of the effect on changes in a stock's return over six months of removing the "uptick rule." Vertical lines indicate a change in the average return of a given magnitude (and therefore its probability). The impact is large and surely economically significant as it would with high probability eliminate a major proportion of the average increase in stock returns. The small change between SEC analysis and the improved analysis we performed makes the tail of the distribution to the right of the zero have less than 5 %, and thus the impact of the uptick rule repeal is statistically significant at the 95 % level. The original SEC paper dismissed the importance of the uptick rule based upon the lack of statistical significance in their analysis—making the classic error of confounding significance and statistical significance, and also making an error in the evaluation of statistical significance

The autocratic framework can be directly investigated by studies of individual molecular mechanisms and has been the starting point for discussions of biological processes. But a broader understanding of regulatory mechanisms is needed that incorporates essential features of both extreme views. The democratic framework relies on mutual regulation, which tends toward a self-consistent gene expression state that is stable in the face of fluctuations. In other words, this view has its roots in the conceptual understanding of stability and homeostasis of cell types. The democratic view has only recently gained empirical support, perhaps because its characterization involves studies of genome-wide dynamical processes.

What framework should be used to study collective state control? The difficulty is that for individual gene effects, individual transcription levels are important. For attractors, collective dynamics of the transcriptome within a cell type, rather than specific gene expression signatures, characterize cell behavior (not just cell type differences). What is needed are control coefficients that measure change in collective states relative to archetypes [4], and in relation to individual gene transcription level

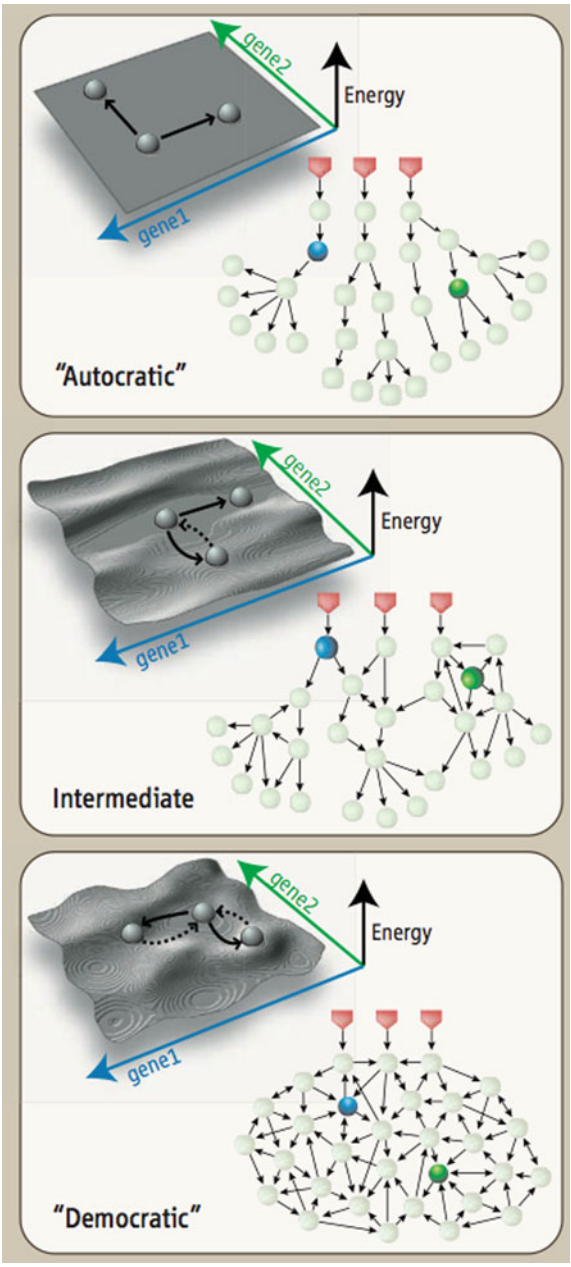


Fig. 4 Transcription regulatory architecture. In autocratic regulatory networks (*top*), individual master regulator genes (*pointed squares*) are stimulated by external signals and control many other genes (*circles*). As shown by the energy landscape, the transcriptional states (*spheres*) may have no preferences (*black arrows* represent changes in expression of genes 1 and 2). In democratic networks (*bottom*), all genes act as mutual regulators. A few specific gene expression patterns become stable, shown as basins of attraction (cell types) in the landscape. Once a cell reaches one of these states, changing the expression of one gene is unlikely to switch the cell type (*black arrows*). Intermediate networks (*middle*) have mutual regulation, but certain genes (*blue circle*) are major controllers. For details see [4]

changes. Relating the variation of small sets of gene expression values to deviation or conformity to archetypes can provide a framework to study the interplay of attractors and master regulators. Such observations, best taken from unaveraged data, should

identify the dispersal and convergence of cells near an attractor, and the mechanisms of homeostatic control. Using multiple archetypes also should enable the study of cell fate trajectories.

2.2 Empirical Multiscale Networks of Cellular Regulation

In a eukaryotic organism such as the mouse, the complete transcriptional network contains ≈15,000 genes and up to 225 million regulatory relationships between pairs of genes. Determining all of these relationships is currently intractable using traditional experimental techniques, and, thus, a comprehensive description of the entire mouse transcriptional network is elusive. Alternatively, one can apply the limited amount of experimental data to determine the entire transcriptional network at a less detailed, higher level. This is analogous to considering a map of the world resolved to the kilometer rather than to the millimeter. We derived from mouse microarray data several high-scale transcriptional networks by determining the mutual effective regulatory influences of large modules of genes (see Fig. 5). In particular, global transcriptional networks containing 12–72 modules are derived, and analysis of these multiscale networks reveals properties of the transcriptional network that are universal at all scales (e.g., maintenance of homeostasis) and properties that vary as a function of scale (e.g., the fractions of module pairs that exert mutual regulation). In addition, we describe how cellular functions associated with large modules (those containing many genes) are composed of more specific functions associated with smaller modules.

Notice that a multiscale approach is conceptually essential given the organization of living systems into structures at many scales, and is critical given the staggering challenge of obtaining a complete description of pairwise gene interactions. Still, in view of the complexity of biological function, there is a large amount of information that arises from a multiscale analysis. In [5] we performed an analysis that can be considered as foundational to the development of many other results. It was a high- throughput analysis methodology analogous to high-throughput experimental methods of genome sequencing or gene expression data collection; through our approach, a seemingly overwhelming amount of data is generated by analysis of the large number of regulatory interactions of modules across multiple scales. Our analysis of these results has been correspondingly multiscale. First, we identified global principles, such as the many facets of homeostasis and universality of regulatory effects at larger scales. Second, we found new patterns of multiscale organization, such as the dichotomous distributions of the number of regulatory inputs and outputs at various scales, theincreased target specificity and speed of

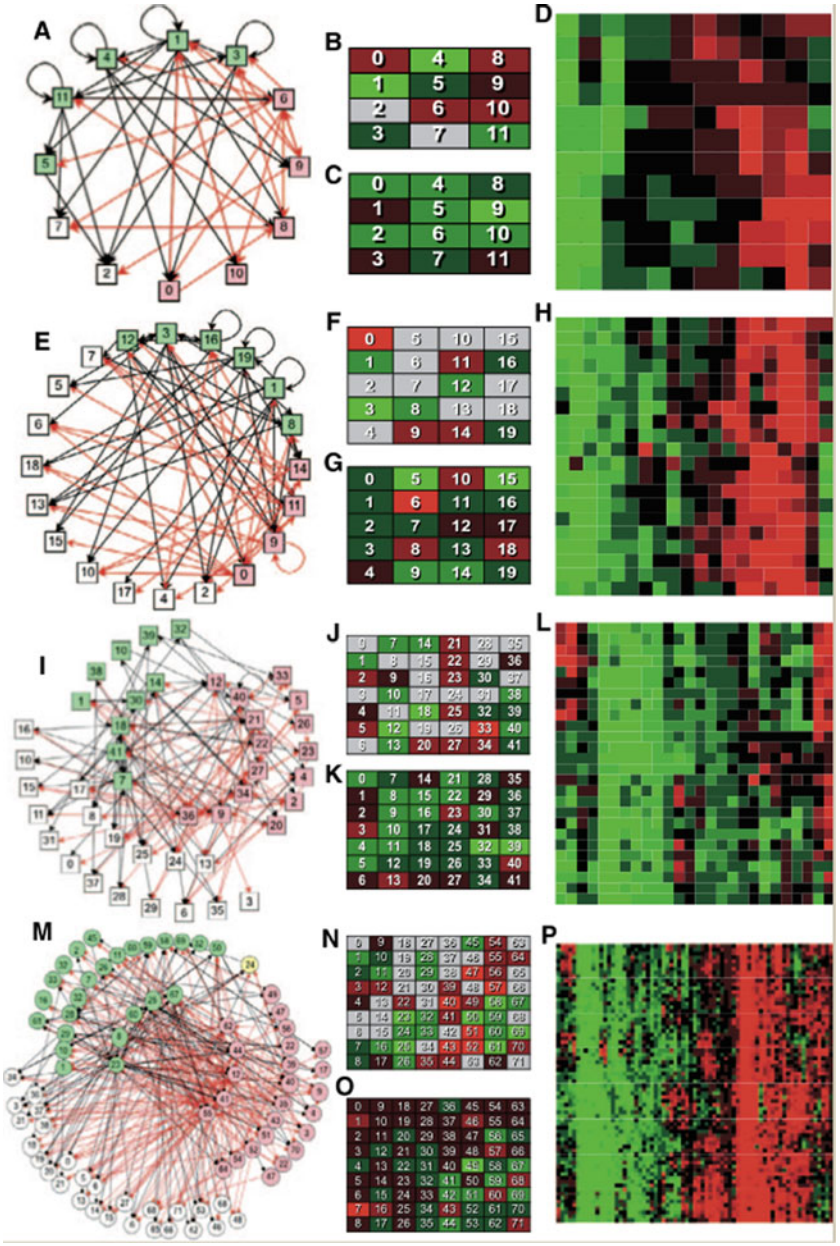


Fig. 5 Effective Regulatory Influences at Many Scales (*Top* to *Bottom*) over a 1.5-h Interval. (A, E, I, M): Regulatory influence network. (B-C, F-G, J-K, N-O): Average magnitudes of outputs (B, F, J, N) and inputs (C, G, K, O) of modules arranged in self-organizing map array order. Stronger inhibition, activation, or neither are indicated by brighter *red*, *green*, or *gray*, respectively. (D, H, L, P) Complete unthresholded regulatory transition matrices. Rows are sorted by similarity in functional inputs, and columns are sorted by similarity in functional outputs. For details see [5]

regulation at finer scales, and the aggregation of sub-module functions into collective larger-scale functions. Last, we provided a detailed discussion of many specific regulatory relationships. The diversity of analysis points the way to many new lines of investigation, in particular experimentally testable hypotheses at large scales of cellular organization.

2.3 Dynamics of Cellular Level Function

A major open question of systems biology is how genetic and molecular components interact to create phenotypes at the cellular level. Although much recent effort has been dedicated to inferring effective regulatory influences within small networks of genes, the power of microarray bioinformatics has yet to be used to determine functional influences at the cellular level. In all cases of data-driven parameter estimation, the number of model parameters estimable from a set of data is strictly limited by the size of that set. Rather than infer parameters describing the detailed interactions of just a few genes, we chose a larger-scale investigation so that the cumulative effects of all gene interactions could be analyzed to identify the dynamics of cellular-level function. By aggregating genes into large groups with related behaviors (megamodules), we were able to determine the effective aggregate regulatory influences among 12 major gene groups in murine B lymphocytes over a variety of time steps. Intriguing observations about the behavior of cells at this high level of abstraction include: (i) a medium-term critical global transcriptional dependence on ATP-generating genes in the mitochondria, (ii) a longer-term dependence on glycolytic genes, (iii) the dual role of chromatin-reorganizing genes in transcriptional activation and repression, (iv) homeostasis-favoring influences, (v) the indication that, as a group, G protein-mediated signals are not concentration-dependent in their influence on target gene expression, and (vi) short-term-activating long-term-repressing behavior of the cell-cycle system that reflects its oscillatory behavior (Fig. 6).

This model is unique, because it is comprehensive in describing all major cellular regulatory influences that occur over a 1.5-h time step, and because it sets forth dozens of experimentally falsifiable hypotheses. Although this analysis has been performed at the megamodule level, this technique can conceptually be used to infer all transcriptional interactions, because its resolution is limited only by the number of time-step experiments. Practically, however, to infer strictly linear effects of every gene, given the levels of experimental noise in microarray data would require $\approx$90,000 observations in mammalian systems or 18,000 observations in yeast. While higher levels of detail can be achieved as described in the previous section, the analysis of the largest scale cell behavior provides key insights into cell function and regulation.

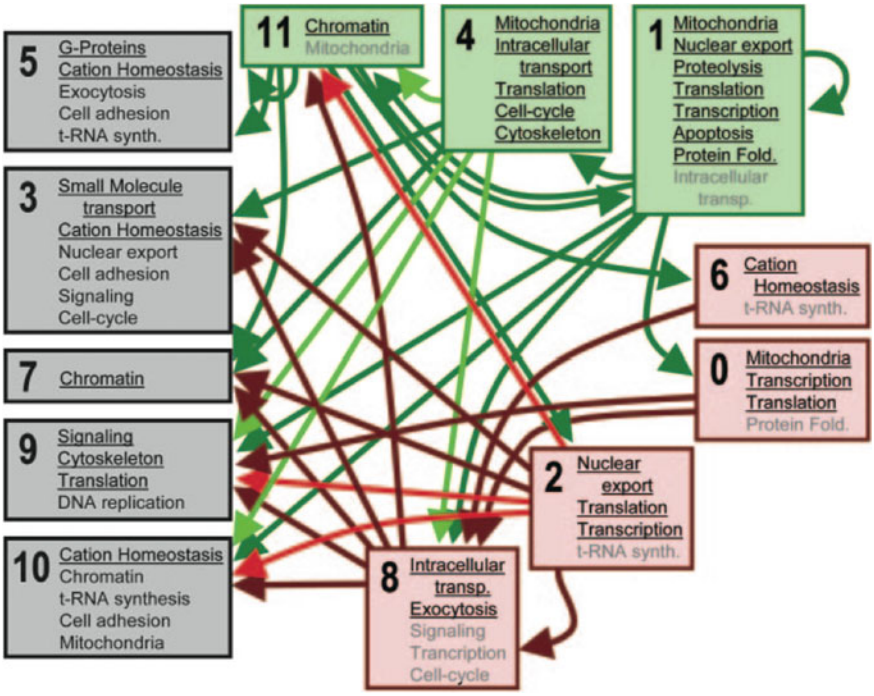


Fig. 6 A global comprehensive cellular influence network for the 1.5-h transition. Each gene group is represented by its number, along with those ontology labels that are overrepresented in that gene group. *Underlined labels* are given when an entire function category is overrepresented, and grayed labels when several individually related functions are each overrepresented, but the category is not. A *green arrow* connecting one gene group to another indicates activation over 1.5 h, *red arrows* indicates repression, with *brighter arrows* indicating stronger effects. For details see [6]

References

1. Harmon, D., Stacey, B., Bar-Yam, Y., Bar-Yam, Y.: Networks of economic market interdependence and systemic risk. arXiv:1011.3707v2 (2010)
2. Lagi, M., Bar-Yam, Y., Bertrand, K. Z., Bar-Yam, Y.: The food crises: a quantitative model of food prices including speculators and ethanol conversion. arXiv:1109.4859 (2011)
3. Harmon, D., Bar-Yam, Y.: Technical Report on SEC Uptick Repeal Pilot. NECSI Technical Reports 2008–2011 http://www.necsi.edu/research/UptickTechReport.pdf (2008)
4. Bar-Yam, Y., Harmon, D., de Bivort, B.: Attractors and democratic dynamics. Science **323**, 5917 (2009)
5. de Bivort, B., Huang, S., Bar-Yam, Y.: Empirical multiscale networks of cellular regulation. PLoS Comput. Biol. **3**, e207 (2007)
6. de Bivort, B., Huang, S., Bar-Yam, Y.: Dynamics of cellular-level function and regulation derived from Murine expression array data. PNAS **101**, 17687 (2004)

Hidden Complexity of Evolutionary Dynamics: Analysis

Ivan Zelinka, Lenka Skanderova, Petr Saloun, Roman Senkerik and Michal Pluhacek

Abstract This chapter presents a method for visualization of the dynamics of evolutionary algorithms in the form of complex networks and is continuation of our previous research. The analogy between individuals of populations in an arbitrary evolutionary algorithm and vertices of a complex network is mentioned, as well as between edges in a complex network and communication between individuals in a population. Visualization of various attributes of network based on differential algorithm is presented here.

Keywords Differential evolution · Complex network · Evolutionary dynamics

1 Introduction

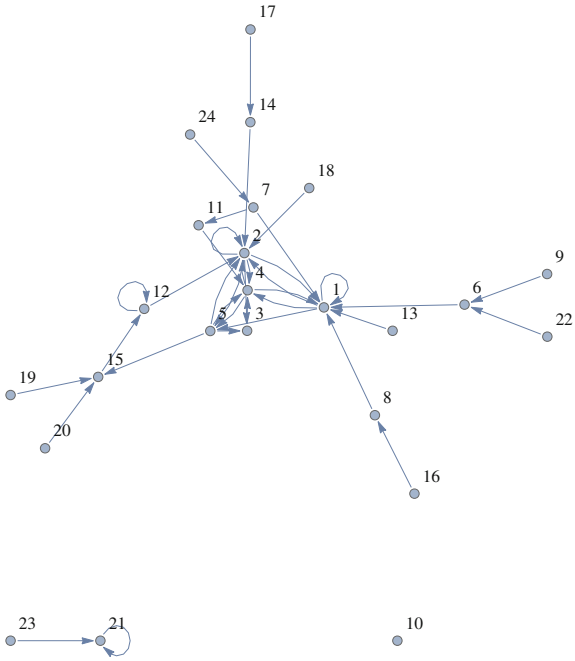
Our previous research in the [1–3] shows that evolutionary algorithm dynamics can convert to the complex network that exhibits similar or the same structural properties as social networks, like Facebook etc. In this chapter we try to propose our interpretations of complex networks analysis tools on complex networks that are given by evolutionary dynamics, see [1–3]. Large-scale networks, exhibiting complex patterns of interaction amongst vertices exist in both nature and in man-made systems (i.e., communication networks, genetic pathways, ecological or economical networks, social networks, networks of various scientific collaboration, Internet,

I. Zelinka (✉) · L. Skanderova · P. Saloun
VSB-Technical University of Ostrava, 17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
e-mail: ivan.zelinka@vsb.cz

R. Senkerik · M. Pluhacek
Faculty of Applied Informatics, Tomas Bata University in Zlin, Zlín, Czech Republic
e-mail: senkerik@fai.utb.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_4,

Fig. 1 Adjacency graph



World Wide Web, power grid etc.). The structure of complex networks thus can be observed in many systems. The word complex networks [4, 5] comes from the fact that they exhibit substantial and non-trivial topological features with patterns of connection between vertices that are neither purely regular nor purely random. Such features amongst other include a heavy tail in the degree distribution, a high clustering coefficient, hierarchical structure. In the case of directed networks, these features also include reciprocity, triad significance profile and other features.

Amongst many studies, two well-known and much studied classes of complex networks are the scale-free networks and small-world networks (see example in Fig. 1), whose discovery and definition are vitally important in the scope of this research. Specific structural features can be observed in both classes i.e. so called power-law degree distributions for the scale-free networks and short path lengths with high clustering for the small-world networks. Research in the field of complex networks has joined together researchers from many fields such as mathematics, physics, biology, chemistry computer science, epidemiology etc. Social network researchers have acquired data for their studies using various methods. In the past these studies were only based on questionnaire data which typically reached the amount of hundreds individuals [6]. In the late 1990s new technologies such as internet and cellular phones enabled researchers to construct large scale networks using emails [7], phone records [8] or web search engines [9].

Many researchers focus on identifying and analyzing community structures in networks growing from web users. Analysis of topological characteristics of the

tripartite hyper graph of queries, users and bookmarks on a large snapshot of web site and on query logs of two large search engines is described in [10]. The extensive analysis of characteristics of large online social network MySpace was published in [11]. Study was oriented to the sociability of users based on relationship, messaging, and group participation, on the demographic characteristics of users with emphasis on correlation with their privacy references and on text analysis with intention to construct language models used by MySpace users. In [12] researchers analyzed 70 large sparse real-world networks and defined network community profile plot which characterizes the "best" profile community. They compare and contrast several methods to approximately optimize metric, based on conductance measure.

Structure analysis of social networks provide large amount of worthy information about interactions between nodes (users) in the network. Understanding the evolution of obtained network can serve more additional information of network behavior in progress. These findings can help to protect such networks from various attacks or can contribute in research of data distribution throughout the network. Several models of network evolution were proposed. In [13] analysis of large social web networks like Flickr and Yahoo is presented. Authors analyzed the network evolution and defined three types of network components on the basis of their activity - singletons, giant components and middle region. Approaches to link prediction of new interactions amongst members in a social network based on measures for analyzing the proximity of nodes were developed in [14]. Theoretical model which explains evolution of communities in social networks where new friends are formed by looking at friends of friends was published in [15]. Group analysis of information in large social networks like Live Journal in relation to its growth and evolution is mentioned in [16], DBLP database is studied in [17]. Authors claim that network cannot be analyzed independently, but need to be studied in the context of other networks.

1.1 Evolutionary Dynamics and Complexity

Motivation of this research is quite simple. As mentioned in the introduction and reported in [3], evolutionary algorithms are capable of hard problem solving. A number of examples on evolutionary algorithms can be easily found. Evolutionary algorithms (EA) use with chaotic systems is done for example in [18] where EAs has been used on local optimization of chaos, [19] for chaos control with use of the multi-objective cost function or in [20] and [21], where evolutionary algorithms have been studied on chaotic landscapes. Slightly different approach with evolutionary algorithms is presented in [22] where selected algorithms were used to synthesize artificial chaotic systems. In [23, 24] EAs has been successfully used for real-time chaos control and in [25] EAs was used for optimization of Chaos Control.

On the other hand, complex networks widely studied across many branches of science are promising and they are intiating a modern interdisciplinary research. Evolutionary algorithms, based on its canonical central dogma (following darwinian ideas) clearly demonstrate intensive interaction amongst individual in the population,

which is, in general, one of the important attributes of complex networks (intensive interaction amongst the vertices).

The main ambition of this chapter is to interpret tools of complex networks analysis on complex networks that are given by evolutionary dynamics. Reason for this is that today various techniques for analysis and control of complex networks exist. If complex network structure is hidden behind EA dynamics, then we believe, that for example above mentioned control techniques could be used to improve dynamics of EAs. The first steps (i.e. conversion of the EA dynamics to the complex network and to CML system) have been done in the [1–3]. Now we propose how standard tools of complex networks analysis can be understood for EAs dynamics purposes.

1.2 Experiment Design

1.2.1 Used Algorithm

For the experiments described here, stochastic optimization algorithms, such as DE [26] have been used in order to get sample data, that were used for visualization of DE complex network (multiple edges are not visualized) and further analysis, see Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. The low number of generations has been set (30) to get network that will be suitable for simple and clear visualization. Analysis has been done in Mathematica version 9. Differential Evolution [26] is a population-based optimization method that works on real-number-coded individuals. For each individual $\mathbf{x}_{i,G}$ in the current generation G, DE generates a new trial individual $\mathbf{x'}_{i,G}$ by adding the weighted difference between two randomly selected individuals $\mathbf{x}_{r1,G}$ and $\mathbf{x}_{r2,G}$ to a randomly selected third individual $\mathbf{x}_{r3,G}$. The resulting individual $\mathbf{x'}_{i,G}$ is crossed-over with the original individual $\mathbf{x}_{i,G}$. The fitness of the resulting individual, referred to as a perturbed vector $\mathbf{u}_{i,G+1}$, is then compared with the fitness of $\mathbf{x}_{i,G}$. If the fitness of $\mathbf{u}_{i,G+1}$ is greater than the fitness of $\mathbf{x}_{i,G}$, then $\mathbf{x}_{i,G}$ is replaced with $\mathbf{u}_{i,G+1}$; otherwise, $\mathbf{x}_{i,G}$ remains in the population as $\mathbf{x}_{i,G+1}$. DE is quite robust, fast, and effective, with a global optimization ability. It does not require the objective function to be differentiable, and it works well even with noisy, epistatic and time-dependent objective functions. Scheme of DE is given as a pseudocode by Eq. 1, see also [3]. DE parameters are in the Table 1.

Fig. 2 Graph partition

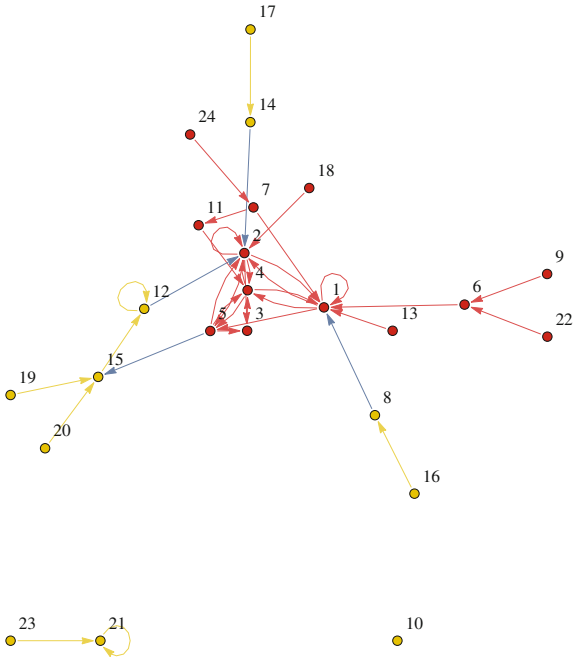


Fig. 3 Minimum cut

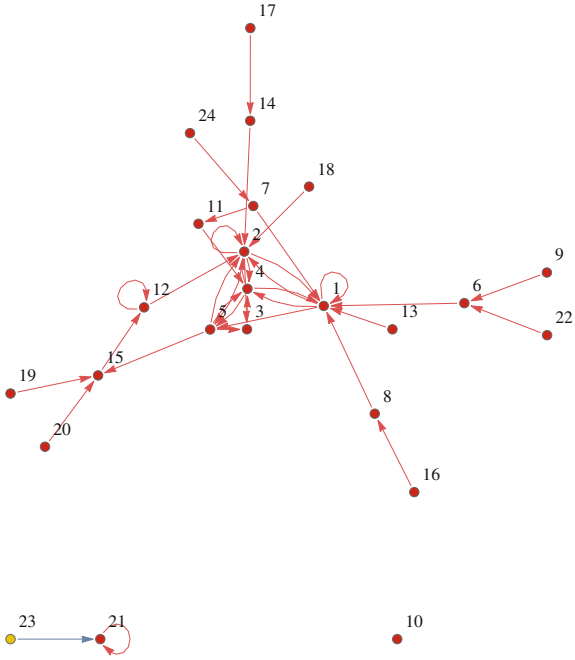


Fig. 4 Degree centrality

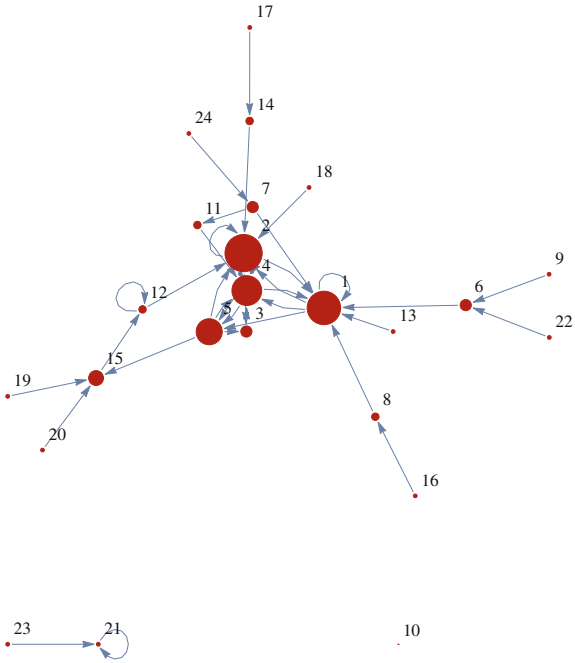
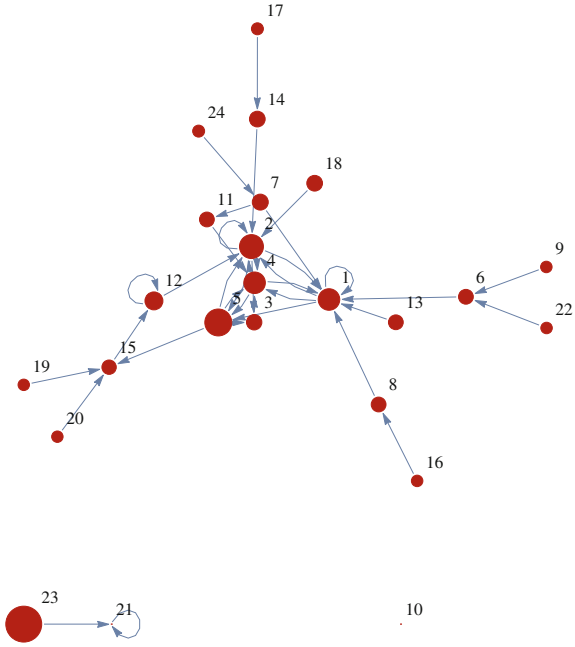


Fig. 5 Closeness centrality



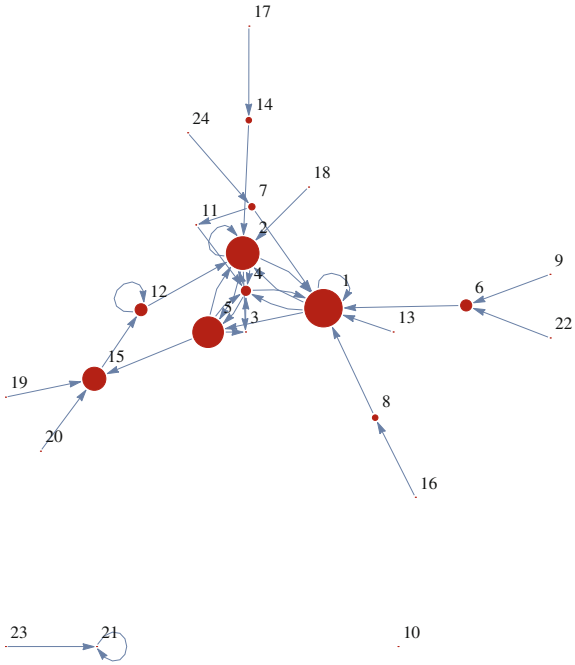


Fig. 6 Betweenness centrality

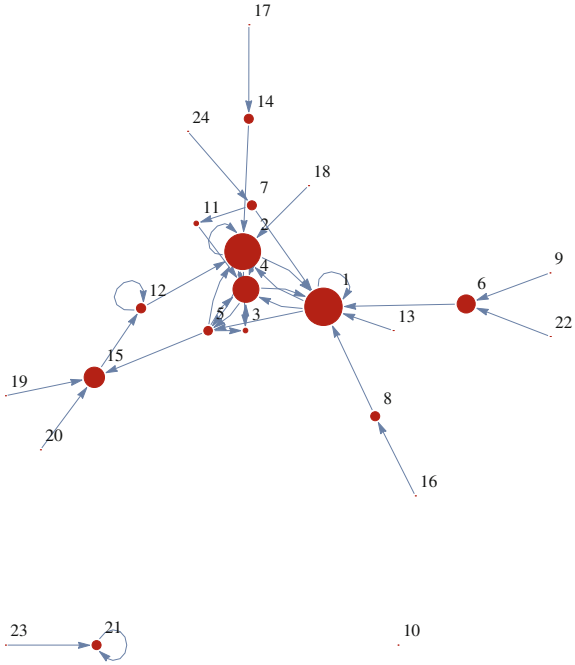


Fig. 7 Page rank centrality

Fig. 8 Mean neighbor degree

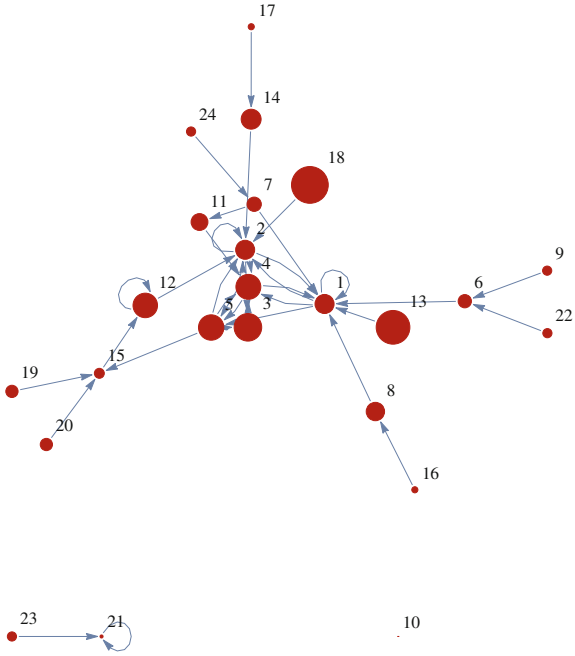


Fig. 9 K core components

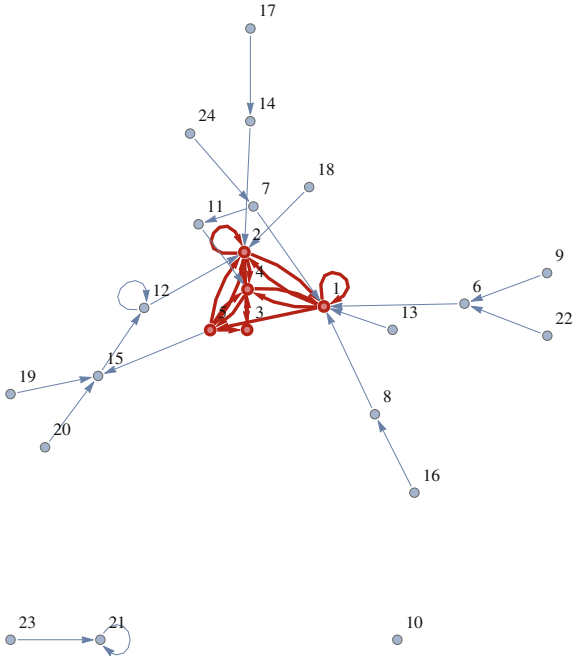


Fig. 10 K plex

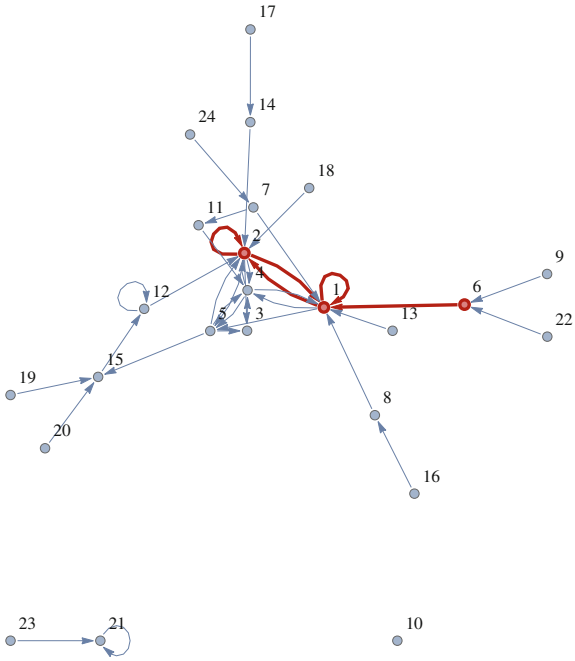


Fig. 11 K club

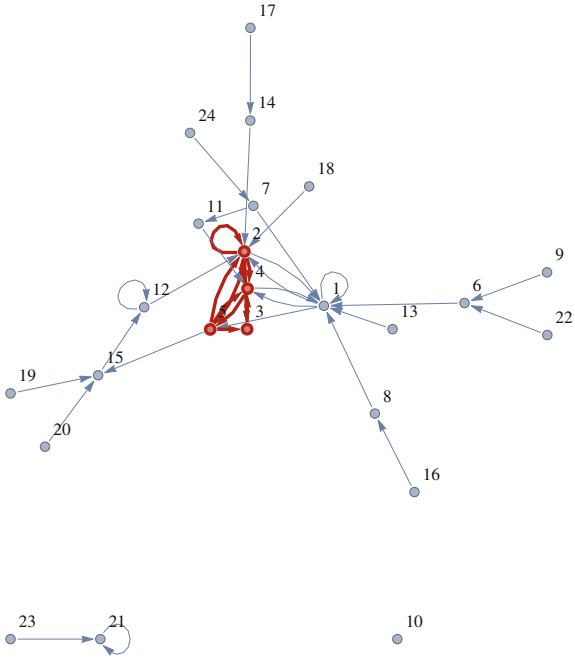


Fig. 12 K clan

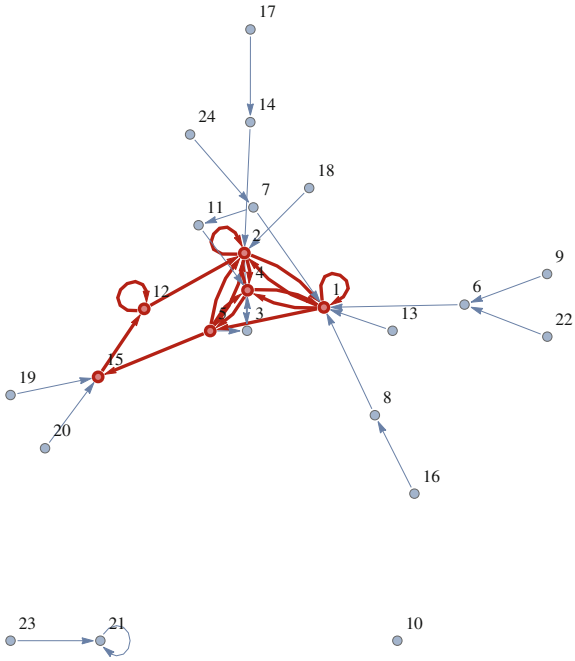


Fig. 13 K clique

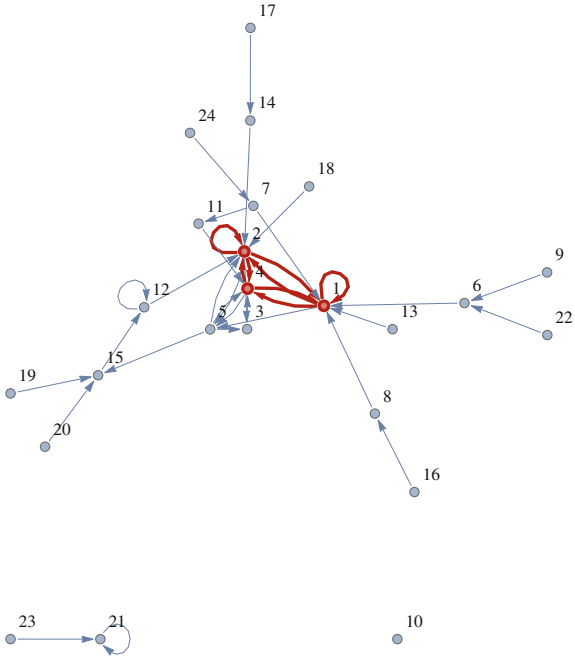


Fig. 14 Clique

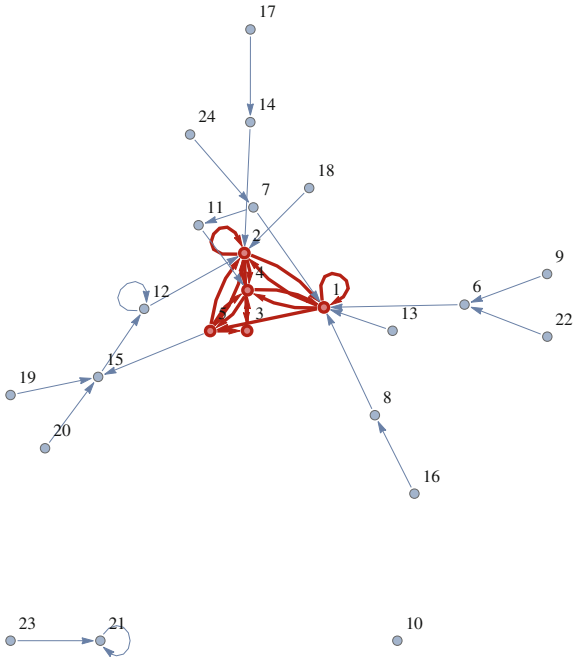


Fig. 15 Community graph

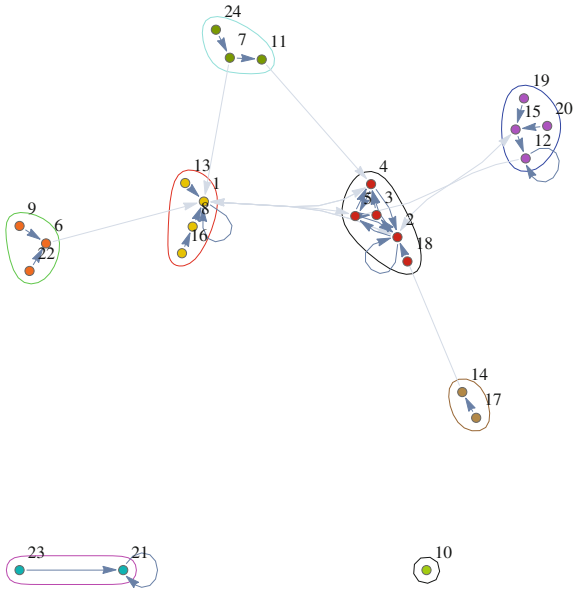


Table 1 DE setting

NP	24
F	0.9
Cr	0.3
Generations	30
Individual length	50

$$
\begin{cases}
1.\text{Input :} D, G_{\max}, NP \geq 4, F \in (0,1+), CR \in [0,1], \text{ and initial bounds :} \mathbf{x}^{(lo)}, \mathbf{x}^{(hi)}. \\
2.\text{Initialize :} \begin{cases} \forall i \leq NP \wedge \forall j \leq D : x_{i,j,G=0} = x_j^{(lo)} + rand_j\,[0,1] \bullet \left(x_j^{(hi)} - x_j^{(lo)}\right) \\ i = \{1,2,\ldots,NP\},\ j = \{1,2,\ldots,D\},\ G = 0, rand_j[0,1] \in [0,1] \end{cases} \\
\begin{cases} 3.\text{While } G < G_{\max} \\ \forall i \leq NP \begin{cases} 4.\text{Mutate and recombine :} \\ 4.1 r_1, r_2, r_3 \in \{1,2,\ldots,NP\}, \text{randomly selected, except :} r_1 \neq r_2 \neq r_3 \neq i \\ 4.2 j_{rand} \in \{1,2,\ldots,D\}, \text{randomly selected once each } i \\ 4.3 \forall j \leq D, u_{j,i,G+1} = \begin{cases} x_{j,r_3,G} + F \cdot (x_{j,r_1,G} - x_{j,r_2,G}) \\ \text{if}(rand_j[0,1] < CR \vee j = j_{rand}) \\ x_{j,i,G} \text{otherwise} \end{cases} \\ 5.\text{Select} \\ \mathbf{x}_{i,G+1} = \begin{cases} \mathbf{u}_{i,G+1} \text{ if } f(\mathbf{u}_{i,G+1}) \leq f(\mathbf{x}_{i,G}) \\ \mathbf{x}_{i,G} \text{ otherwise} \end{cases} \end{cases} \\ G = G + 1 \end{cases}
\end{cases}
\tag{1}
$$

1.2.2 Selected Test Functions and Its Dimensionality

The test function applied in this experimentation was selected from the test bed of 17 test functions—Schwefels function (2). Evolution was searching for global extreme in 50 dimensions, i.e. individual has 50 parameters. Dimension is in the formula (2) represented by variable D, so as one can see, it is easy to calculate selected functions for an arbitrary dimension. Schwefels function has been selected due to their various complexity and mainly for the fact that these functions are widely used by researchers working with evolutionary algorithms. Another reason was that speed of convergence and thus evolutionary dynamics itself is different on that function, compared to simple test functions.

$$\sum_{i=1}^{D} -x_i \sin(\sqrt{|x_i|}) \tag{2}$$

1.2.3 Data for Complex Network Visualization

The most critical point of this research and related simulations was to which data and relations should be selected and consequently visualized. Based on investigated

algorithms, we believe that there is no universal approach, but rather a personal one, based on the knowledge of algorithm principle. In the case of the DE, e.g. DERand1Bin in which each individual is selected in each generation to be a parent. Thus in DE, we have recorded only those individuals-parents, that has been replaced by better offspring (like vertex with added connections). In the DE class of algorithms we have omitted the philosophy that a bad parent is replaced by a better offspring, but accepted philosophical interpretation, that individual (worse parent) is moving to the better position (better offspring in original DE philosophy). Thus no vertex (individual) has to be either destroyed or replaced in the philosophical point of view. If, for example, DERand1Bin has a parent replaced by offspring, then it was considered as an activation (new additional links, edges) of vertex-worse parent from three another vertices (randomly selected individuals, see [1–3, 26]).

1.2.4 Visualization Methods

Experimental data are visualized in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15. Interactions between individuals in the population during entire evolution are described like edges, vertices in graph are individuals that are activated by other individuals, incrementally from generation to generation.

1.3 Interpretation of the Example Network

As reported above DE algorithms were set according to Table 1 and have been tested on Schwefel's 2 test function (to reveal its complex networks dynamics) with constant level of test function dimensionality (i.e. individual length). All data have been processed graphically. Visual emergence of complex network structure behind evolutionary dynamics depends on many factors. However some special versions of used algorithms did not show complex network structure despite the fact that the number of generations was quite large, see [1–3]. The main tools of *Mathematica* software were used on basic analysis and are proposed here. We would like to remark that graphs on following figures depend usually on command setting that generate graph. We have used basic setting of used Mathematic commands. Visualized and analyzed graphs have multiple edges that can be understand like weight of single edge. Attributes of proposed analysis are represented by subgraph colors and vertice sizes in graphs. Our proposed interpretation, based on terms and command from Wolfram *Mathematica* used for all of our experiments is following:

1. **Adjacency graph**, see Fig. 1.
   Meaning: Graph with vertices and oriented edges.
   Interpretation: Visualization of evolutionary dynamics in the form of so called graph. Each vertex represent one individual in the population and each edge (oriented of course) represent successful offspring creation (i.e. fitness

improvement of active parent in this philosophy) between parents connected by that edge.

2. **Graph partition**, see Fig. 2.
   Meaning: Graph partition finds a partition of vertices such that the number of edges having endpoints in different parts is minimized. For a weighted graph, graph partition finds a partition such that the sum of edge weights for edges having endpoints in different parts is minimized.
   Interpretation: Individuals in population are separated into "groups" according to their interactions with another individuals, based on their success in active individual fitness improvements. "Endpoints" can be understood like successful participation of selected individuals in active individual fitness. On Fig. 2 is partition visualized by colors. This analysis gives view on population structure and shows the set of individuals that got or donate oriented edges (support from / to) the same group of individuals. Based on number of connections or weights (if multiple edges are understood like integer weights) of edge, it can be analyzed what part of population was the most important in the evolutionary dynamics for given case.
3. **Minimum cut**, see Fig. 3.
   Meaning: A minimum k-cut of a graph g is a partition of vertices of g into k disjoint subsets with the smallest number of edges between them. For weighted graphs, minimal cut gives a partition c_1, c_2,... with the smallest sum of edge weights possible between the sets c_i.
   Interpretation: It is quite similar to previous item, population is separated into disjoint subsets with the smallest number of interactions between individuals. The expression "disjoint subsets with the smallest number of edges between them" say that minimal cut gives set of individuals that has the lowest impact on evolutionary dynamics. It allows to study impact of evolutionary dynamics on population structure from interaction point of view.
4. **Degree centrality**, see Fig. 4.
   Meaning: Degree centrality of g gives a list of vertex degrees for the vertices in the underlying simple graph of g. Degree centrality will give high centralities to vertices that have high vertex degrees. The vertex degree for a vertex v is the number of edges incident to v. For a directed graph, the in-degree is the number of incoming edges and the out-degree is the number of outgoing edges. For an undirected graph, in-degree and out-degree coincide.
   Interpretation: Degree centrality shows how many in-coming (support from individuals) or out-coming (support to individuals) edges vertex - individual under study has. This quantity can be related to progress of the evolutionary search and used to made conclusion of what set of individuals has maximally contribute to that. On Fig. 4 are individuals sized according to that degree.
5. **Closeness centrality**, see Fig. 5.
   Meaning: Closeness centrality will give high centralities to vertices that are at a short average distance to every other reachable vertex. ClosenessCentrality for a graph is given by $\left(\frac{1}{l_1}, \frac{1}{l_2}, \ldots\right)$, where l_i is the average distance from vertex i

to all other vertices connected to i. If d is the distance matrix, then the average distance l_i from vertex i to all connected vertices is given by $\frac{\sum_j d_{i,j}}{k}$, where the sum is taken over all finite $d_{i,j}$ and k is the number of vertices connected to i. The closeness centrality for isolated vertices is taken to be zero.
Interpretation: Closeness centrality shows which of individuals has the short average distance to every of the other individuals. In fact such individual has shortest path to the most other individuals and thus information about its fitness and its genome is easily spread throughout the population (graph vertices), see Fig. 5.

6. **Betweenness centrality**, see Fig. 6.
Meaning: Betweenness centrality will give high centralities to vertices that are on many shortest paths of other vertex pairs. Betweenness centrality for a vertex i in a connected graph is given by $\sum_{s,t \in v \wedge s \neq i \wedge t \neq i} \frac{n^i_{s,t}}{n_{s,t}}$, where $n_{s,t}$ is the number of shortest paths from s to t and $n^i_{s,t}$ i is the number of shortest paths from s to t passing through i. The ratio $\frac{n^i_{s,t}}{n_{s,t}}$ is taken to be zero when there is no path from s to t.
Interpretation: Betweenness centrality shows individuals with high centrality, that are on the shortest paths-connections with other individuals. In fact, it shows individuals, through which is information about fitness (or influence of the most successful individuals) spread the most intensively through the population.
7. **Page rank centrality**, see Fig. 7.
Meaning: Page rank centrality gives a list of centralities that are solutions to $c = \alpha a^T.d.c + \beta$, where a is the adjacency matrix of g and d is the diagonal matrix consisting of $1/\max\left(1, d_i^{\text{out}}\right)$, where d_i^{out} is the out-degree of the i^{th} vertex.
Interpretation: Page rank centrality is another view on usability of given individual and can be interpreted like higher page rank centrality - higher and more frequent use of individual in evolutionary search and that has very intensive impact on population development.
8. **Mean neighbor degree**, see Fig. 8.
Meaning: The mean neighbor degree is also known as the average neighbor degree. The mean neighbor degree of the vertex v is the mean of vertex degrees of neighbors of v. For weighted graphs, the mean neighbor degree of the vertex v is given by $\frac{\sum_u d_u w_{\text{uv}}}{s}$ over all neighbors u of v with edge weight w_{uv} between u and v. Variable d_u is the degree of the vertex u and s is the total of weights w_{uv}.
Interpretation: The mean neighbor degree gives the average neighbor degree of surrounding (in the network connection sense) individuals and can be used like statistical information how population neighbour of individual is formed, based on information that is spread through population. In fact, the mean neighbor degree can give the map over the whole population showing "density" of importance of individuals collections/sets.

9. **K core components**, see Fig. 9.
   Meaning: A k-core component is a maximal weakly connected subgraph in which all vertices have degree at least k and can be formed like a list of components c_1, c_2,..., where each component c_i is given as a list of vertices. For a directed graph g, KCoreComponents gives the k-core components of the underlying undirected simple graph of g.
   Interpretation: A k-core component show a set of individuals with degree at least k, so individuals in the population can be grouped under this "filter" for further analysis.
10. **K-plex**, see Fig. 10.
    Meaning: A k-plex is a maximal set of vertices such that each vertex is adjacent to all except "k" others. For a directed graph, the outgoing edges for each vertex connect to all except "k" others.
    Interpretation: Show individuals that have at least connections with other excluding "k" others. It allows filter out of population individuals for further analysis.
11. **K-club**, see Fig. 11.
    Meaning: A k-club is a maximal set of vertices where the diameter of the corresponding subgraph is at most k.
    Interpretation: Shows a set of individuals, whose interaction (vizualized in graph like edges) is mutually through k individuals. This can be used to study spread of information throughout the population.
12. **K-clan**, see Fig. 12.
    Meaning: A k-clan is a k-clique where the diameter of the corresponding subgraph is at most k.
    Interpretation: Individuals subgraph that has diameter k. This is another view on patters structure formatting in the population.
13. **K clique**, see Fig. 13.
    Meaning: A k-clique is a maximal set of vertices that are at a distance no greater than k from each other.
    Interpretation: A maximal set of individuals that are at a interaction distance no greater than k from each other individuals.
14. **Clique**, see Fig. 14.
    Meaning: A clique is a maximal set of vertices where the corresponding subgraph is a complete graph.
    Interpretation: Limit case of k-clique, a maximal set of individuals where the corresponding subgraph is a complete graph - i.e. all individuals are in connection with other from that sub-graph.
15. **Community**, see Fig. 15.
    Meaning: Community graph plot attempts to draw the vertices grouped into communities.
    Interpretation: Community graph plot showing the individuals grouped into communities. Communities (with border are individuals that communicate amongst themselves (higher density of edges in community, multi edges are not visualized here, rather than between communities) and community are then

joined by connections that are "one-way" and shows flow of information between communities). This kind of visualization can be interesting also in the case of parallel EAs, where islands of subpopulations are formed.

1.4 Conclusion

In this chapter we have suggested possible interpretation of selected well known tools and terminology from complex networks analysis to the evolutionary algorithms dynamics converted to the complex network structures. The volume of this article is too small to mention and explain all the possible interpretations and tools. This is only a mid-step in our research presented in [1–3], where we proposed all necessary steps joining evolutionary dynamics, complex networks and CML systems.

This chapter is not focused on mathematics itself or on numerical simulations, as in the previous chapters, [1–3], but rather on "philosophical" interpretation of proposed terms used in Wolfram *Mathematica* and in all our past and also future experiments. The next step is to use all previous results to create case study and demonstrate usefulness of proposed ideas in this way. All results will be published in further research papers.

Acknowledgments The following two grants are acknowledged for the financial support provided for this research: Grant Agency of the Czech Republic—GACR P103/13/08195S, by the Development of human resources in research and development of latest soft computing methods and their application in practice project, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic, partially supported by Grant of SGS No. SP2013/114, VŠB - Technical University of Ostrava, Czech Republic, and by European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089.

References

1. Zelinka I., Davendra D., Senkerik R., Jasek R., Do Evolutionary Algorithm Dynamics Create Complex Network Structures? Complex Syst. **20**, 127–140 (2011). ISSN 0891–2513
2. Zelinka I., Davendra D., Snasel V., Jasek R., Senkerik R., Oplatkova Z.: Preliminary investigation on relations between complex networks and evolutionary algorithms dynamics. In: CISIM 2010. Poland (2010)
3. Zelinka I., Davendra D., Chadli M., Senkerik R., Dao T.T., Skanderova, L.: Evolutionary dynamics and complex networks. In: Zelinka I, Snasel V and Abraham, A. (eds.) Handbook of Optimization. Springer Series on Intelligent Systems. Springer, Berlin (2012)
4. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. Adv. Phys. **51**, 1079 (2002)
5. Boccaletti, S., et al.: Complex networks: structure and dynamics. Phys. Rep. **424**, 175–308 (2006)
6. Wasserman, S., Faust, K.: Social Network Analysis: Methods and Applications Structural Analysis in the Social Sciences. Cambridge University Press, Cambridge (1994)
7. Eckmann, J.-P., Moses, E., Sergi, D.: Entropy of dialogues creates coherent structures in e-mail traffic. Proc. Natl. Acad. Sci. USA **101**, 14333–14337 (2004)

8. Onnela, J.-P., Saramaki, J., Hyvonen, J., Szabo, G., de Menezes, A.M., Kaski, K., Barabasi, A.-L., Kertesz, J.: Analysis of large-scale weighted network of one-to-one human communication. New J. Phys. **9**(7), 179 (2007)
9. Lee, S.H., Kim, P.-J., Ahn, Y.-Y., Jeong, H.: Googling social interactions: web search engine based social network construction. arXiv: 0710.3268v1 (2008)
10. Krause, B., et. al.: Logsonomy—social information retrieval with logdata. In: Proceedings of the 19th ACM Conference on Hypertext and Hypermedia, pp. 157–166. Pittsburgh, PA, USA (2008)
11. Caverlee, J., Webb, S.: A large-scale study of myspace: observations and implications for online social networks. In: Association for the Advancement of Artificial Intelligence (2008)
12. Leskovec, J., et al.: Statistical properties of community structure in large social and information networks. In: WWW 2008, pp. 695–704. Beijing, China (2008)
13. Kumar, R., Novak, J., Tomkins, A.: Structure and evolution of online social networks. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 611–617. ACM Press, Philadelphia, PA, USA (2006)
14. Liben-Nowell, D., Kleinberg, J.: The link prediction problem for social networks. J. Am. Soc. Inf. Sci. Technol. (Wiley Periodicals) **58**(7), 1019–1031 (2007)
15. Gollapudi, S., Kenthapadi, K., Panigrahy, R.: Threshold phenomena in the evolution of communities in social networks. In: 17th International World Wide Web Conference (WWW2008), Workshop on Social Web Search and Mining (swsM2008) (2008)
16. Backstrom, L., Huttenlocher, D., Kleinberg, J., Lan, X.: Group formation in large social networks: membership, growth, and evolution. In: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 44–54. ACM Press, Philadelphia, PA, USA (2006)
17. Cai, D., Shao, Z., He, X., Yan, X., Han, J.: Mining hidden community in heterogeneous social networks. In: Proceedings of the 3rd International Workshop on Link Discovery, pp. 58–65. ACM Press, Chicago, Illinois (2005)
18. Richter, H., Reinschke, K.J.: Optimization of local control of chaos by an evolutionary algorithm. Physica D **144**, 309–334 (2000)
19. Richter H.: An evolutionary algorithm for controlling chaos: the use of multi-objective fitness functions. In: Guervos M., Panagiotis J.J., Beyer A., Villacanas F.H.G., Schwefel, J.L., Schwefel H.P. (eds.) Parallel Problem Solving from Nature-PPSN VII. Lecture Notes in Computer Science, vol. 2439, pp. 308–317. Springer, Berlin (2002)
20. Richter H., Evolutionary Optimization in Spatio- temporal Fitness Landscapes. Lecture Notes in Computer Science, NUMB 4193, pp. 1–10. Springer (2006). ISSN 0302–9743
21. Richter H.: A study of dynamic severity in chaotic fitness landscapes. Evolutionary computation, 2005. The IEEE Congress, vol. 3, issue 2–5, pp. 2824–2831 (2005)
22. Zelinka, I., Chen, G., Celikovsky, S.: Chaos synthesis by means of evolutionary algorithms. Int. J. Bifurcat. Chaos Univ. Calif. Berkeley USA **18**(4), 911–942 (2008)
23. Zelinka, I.: Real-time deterministic chaos control by means of selected evolutionary algorithms. Eng. Appl. Artif. Intell. (2008). doi:10.1016/j.engappai.2008.07.008
24. Zelinka I.: Investigation on realtime deterministic chaos control by means of evolutionary algorithms. In: 1st IFAC Conference on Analysis and Control of Chaotic Systems, Reims, France (2006)
25. Senkerik R., Zelinka I., Navratil E.: Optimization of feedback control of chaos by evolutionary algorithms. In: 1st IFAC Conference on Analysis and Control of Chaotic Systems, Reims, France (2006)
26. Price, K.: An introduction to differential evolution. In: Corne, D., Dorigo, M., Glover, F. (eds.) New Ideas in Optimization, pp. 79–108. McGraw-Hill, London (1999)

The Brain Equation

Otto E. Rössler

Abstract The brain equation is a solution to the "second survival problem." The latter is called "positional adaptation." It unlike Darwin's first ("metabolic adaptation") is history-independent. As such it is mathematically well posed. The equation applies to all life forms in the cosmos that live in a structured environment in which survival depends on position in space in a short-term fashion. An eusocial version does not exist. The equation solves, in conjunction with the necessarily attached VR machine, the famous NP-complete "decision-type travelling salesman problem" for finite times. The resulting autonomous optimizer with cognition is susceptible to a "function change" in the sense of Bob Rosen which so far is known empirically only from the human brain.

Keywords Brain equation · Second darwinism · Decision-type travelling-salesman problem · Deductive biology · A.I. · Epigenetic personogenetic function change · Robert Rosen · Gödel incompleteness.

1 Introduction

An equation presented in a talk in 1973 and published in the pertinent proceedings [1] carries this name, first given to it as the title of an unpublished 1973 paper.

The explanation of why the equation is not well known seems to lie in the fact that it was derived in an arcane mathematical context—the so-called "travelling salesman problem" which is "NP-complete." The term "Non-Polynomial complete" means that the solution involves an exponential ("nonpolynomial") factor while the word "complete" refers to the fact that very many applied problems belong into the same class. It also means that if one member of the class can be solved in polynomial time, all can. Even an influential physicist, John Archibald Wheeler, gave the problem

O. E. Rössler (✉)
Faculty of Science, University of Tubingen, Auf der Morgenstelle 8,
72076 Tubingen, Federal Republic of Germany
e-mail: oeross00@yahoo.com

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_5,

a stab. Its mathematical depth is illustrated by the fact that the limiting solution is "Gödel complete" in the sense that it represents a fictitious but mathematically existing example of a Gödel-incomplete solution [2], which is very paradoxical and futuristic.

In spite of its arcane origin, the new "travelling-salesman-with-alarmclocks problem" [1], or synonymously "decision-type travelling salesman problem" (as the authors of the influential book of ref. [3] would later independently call it), is of a maximum applied interest. And so is its solution based on the brain equation. The brain equation enables intelligent autonomous robots and spacecraft to be built. More importantly, it allows for an understanding of the difference between autistic intelligent behavior on the one hand and person-controlled intelligent behavior on the other, no matter whether shown by intelligent machines or living beings. The computer "HAL" in the "Space Odyssey" movie of 1968 illustrates the point. In biology, the "second travelling-salesman problem" (as it can be called for short) complements Darwinism by a purely mathematical sister discipline called "deductive biology" [4]. In other words, the brain equation is a mathematician's dream if it is correct.

The equation, condensed into the short Eq. (13) of ref. [1], comprises several antecedent equations there. It was re-published in ref. [5]. It will not be repeated in the following (unless the reader jumps to the Appendix). Nevertheless all features of the brain equation will be attempted to be made transparent in the following.

2 What the Brain Equation Cannot Do

The brain equation, while representing desires—and in conjunction with the attached universal simulator (or VR machine) presumably all possible thoughts—, cannot describe or even predict the existence of color and other qualia. The latter miraculously step in when the biologically implemented brain equation says "force number N" has this or that strength. For numbers cannot refer to themselves by definition. So the brain equation—like all of science—belongs to the "Hades world" of Greek mythology which obeyed the same laws with its "shadows" as the upper world does but lacked color and "blood" and everything that makes life valuable. Nonetheless the brain equation is maximally efficient. There can be little doubt that, if real brains are miraculously endowed with color and a Now as we cannot deny, an analogous "completion" is granted to the brain equation when it is implemented in a particular hardware.

3 Deductive Darwinism: Part 2

Part one—ordinary Darwinism—describes as is well known the necessary endogenous changes (or changeability) of a population of living systems in case the physical environment changes slowly over time as unavoidable. The theory of the

"brain in the genome" (Michael Conrad's term, personal communication 1979), or synonymously "recursive evolution" [6] which goes back to John Holland and John von Neumann, pays tribute to this fact. This theory is much more complicated and harder to solve than the travelling-salesman-with-alarmclocks problem because it cannot help include unpredictable historical accidents. This fact makes the mathematics of the traditional "temporal Darwinism" far more difficult than that of the new "spatial Darwinism" with its brain equation. Temporal Darwinism is progressing quite slowly as is well known despite the large number of new empirical data that come in every week on many levels.

Traditional temporal Darwinism is non-predictive in very many respects as mentioned. George Kampis (personal communication1991) insists that the shape of an elephant represents a non-predictable biological phenomenon, notwithstanding its intrinsic beauty. The same holds true for the human body and its beauty. The brain equation, by contrast, is thoroughly predictable because the travelling-salesman-with-alarmclocks problem is the same in all biologies, terrestrial and Jovian (the latter likely based on BNBN-chains and liquid ammonia rather than on carbon-chains and liquid water [7]; similarly Georg Graf in a 2010 blog). "Mantis shrimps," mollusks and vertebrates all implement the same brain equation while using different structural means, a fact which automatically makes for a comparative discipline of its own. Only the class of eusocial animals appears to lack a unique brain equation. There is much room for surprises.

The reason for the "divide" between chemical and spatial Darwinism lies in a simple observation. If survival depends acutely on position in space and/or time, the genetic information supplied by the genome is provably incapable of securing survival. The temporal subcase—regarding creatures living in the tidal zone—was actually the trigger which enabled this qualitative difference to be spotted in a discussion with Konrad Lorenz in 1966, as he immediately acknowledged. Darwin's metabolic adaptation is intrinsically unable to foster survival in such short-term cases. Genetically unpredictable spatiotemporal information needs to be recruited here. The resulting Second Darwinism of "positional adaptation" is—unlike the First Darwinism of "metabolic adaptation"—fully predictive. This is because it represents a well-posed mathematical optimization problem as mentioned—the travelling-salesman-with-alarmclocks problem (or synonymously the decision-type travelling salesman problem).

Eric Charnov [8] came very close to the same insight in his theory of "optimal foraging." This second Darwinism unlike the first turns out to be totally deductive. A branch from an arcane mathematical discipline called NP-completeness theory is suddenly in charge.

4 Brief Summary of the Equation

The brain equation when physically embodied yields only a suboptimal—a "local"—solution to the travelling-salesman-with-alarmclocks problem of mathematics. Picture in your mind a landscape with many filling stations of several colors

(yellow, green, red, say) distributed at random, and then picture an autonomous vehicle equipped with equally many on-board fuel tanks of matching color, each requiring to be filled-up independently. No onboard fuel tank of any color must go empty before the next filling station of the same color is reached.

This mathematical optimality problem is virtually trivial if one tank filling carries much farther than the mean distance between filling stations of the same color. Lorenz describes the corresponding series of adaptations in his book, originally titled "The Flip-side of the Mirror" [9]. Yet the same problem becomes insoluble if the two numbers approach each other. As the two numbers are allowed to get closer, an intermediary threshold is reached when a "local solution" no longer suffices and instead a "supra-local solution," chosen as the best among several simulationally calculated possible paths, becomes mandatory. This is the threshold from "direction optimization" towards "path optimization."

The brain equation yields the locally optimal direction, obtained by one's using only the locally available information about the closest-lying source of every color and the momentary filling status of each colored fuel tank. It is only an autonomous "direction optimizer." It is a compromise and as such appears not to deserve much interest at first sight. This local solution—determined from locally evaluated directional weights—is described by the brain equation which proves maximally valuable as we shall see.

The brain equation is a potential-type (more properly speaking: gradient type) equation yielding solutions orthogonal to those of a Newtonian potential equation. It involves 2N additive, nonlinear, positive and negative (that is, attractive and repulsive), position-in-space dependent (as well as time-dependent) sub-potentials—two for each color, one positive, one negative. Some of the explicit nonlinear additive functions that enter can approach infinity—for example, when one fuel tank is almost empty and the closest source of the same color can only be reached by a one hundred percent direct "collision course."

The explicitly written-down equation (Eq. 13 of [1]) contains several idealizations compared to the real-life situation of a human car-driver seeking a specific type of repair shop or filling station: no reserve fuel tanks, equal filling volumes, instantaneous fill-ups when reaching a matching filling station, equal depletion rates of all fuel tanks, and constant travelling speed. However, these assumptions do not make for an appreciable difference as far as the mathematical difficulty of the problem is concerned.

Even though the brain equation only yields a suboptimal, "local" solution as we saw, hardware-wise addition of a "universal simulator" (or synonymously Virtual-Reality—VR—machine or cognitive-map system [10]) renders it effectively optimal in the combination. While the brain equation remains in charge in this combination, other path optimizers (being more chess-like) appear eligible at first sight, too. However, if simulation takes time in proportion to its complexity as unavoidable, then the brain equation is the only remaining alternative—as a "first shot" emergency routine. Therefore, the brain equation turns out to be fundamental. If the environment contains contingencies (not considered so far), then the short-term memory implicit in the universal simulator needs to be complemented by long-term storage devices [1].

By the way, the amount of hardware required for the simulator is bound to approach infinity for the limiting problem of an infinitely long survival. The fictitious pertinent solution then is "Gödel-complete," cf. [2].

It is now almost possible to write the brain equation down explicitly [1, 5]: it describes a directional weight given to every direction in space. This weight has additive contributions from every color, i.e. from every color-specific sub-potential contributing to the attractiveness or repulsiveness of the direction in question. This is because every local direction, if followed, contributes to coming closer or less close to the nearest source of each color (filling station). This weight is either positive or negative-in-the-opposite-direction, or else both weights apply simultaneously. The positive and the negative space dependencies are hereby different (circle-of-Thales vs. cardioid shape). The time dependencies are simple nonlinear functions of the distances and the times left for the respective fuel tank to run out. Moreover, there is a multiplicative global term called "friskiness" present in the equation which is put suppressed (goes to unity) in case there are strong negative contributions. This is all. The resulting equation is a sum of 2N nonlinear terms, partially weighted by the mentioned multiplicative factor greater than unity. The generated directional weights can approach negative or positive infinity.

5 Remark on Social Animals

While the brains of eusocial animals (like ants and naked moles) are not bound to use the brain equation as mentioned, social animals are not exempt from it. In social animals, an obligatory sub-potential of their brain equation is "bonding." Hereby a "mother car" tries to aid a "child car" or "partner car" by being rewarded (increased in its sum functional) by a displayed high level of the parner's sum functional. In vertebrates, the bonding potential is controlled by the hormone oxytocin as is well known empirically (compare Ernst Fehr's game-theoretic studies with human volunteers [11]). The hormone's release is triggered by, among other things, the partner's signaled momentary state of well-being (the value of his or her sum-potential called "happiness"). The book "Neosentience" [12] offers some further details.

6 Discussion

The brain equation has an unusual name. It is not an inductive result, as most findings in science are, but a deductive one. It is "top-down," not "bottom-up." It shares this feature with Darwinism. It belongs into "evolutionary mathematical biology," part of which is deductive as we saw.

The brain equation got largely neglected for four decades despite a few follow-up publications written to draw attention to it (like [13]). This fact is not astounding as the related paradigm of classical Darwinism can teach: the incredible intelligence

hidden in the genome [6] goes unappreciated as a both theoretical and practical treasure trove as mentioned.

At this point it is perhaps worth emphasizing the practical and economic advantages of the brain equation—for example, regarding intelligent robots as needed in the automated care for the elderly in youth-deprived countries. But the brain equation is also not aloof from being applied to real brains. Its complexity after all touches on Gödel incompleteness as we saw. Most importantly, the brain equation, although deterministic of a hyperchaotic type, like many nonlinear systems admits more than one qualitatively different functional regime. More specifically, it can be used not only in the usual "autistic mode" (since its design yields a deterministic—that is, by definition autistic—force field). It can also function in a non-autistic, "person-type" mode. This mathematical implication is an example of the "principle of function change" described by Robert Rosen (cf. his book "Anticipatory Systems" [14]).

This unexpected dynamical implication makes the brain equation relevant for human life not just on the level of industrial gadgets. It teaches us something about ourselves as persons. Let me dwell a little on this in my opinion important point. Two coupled brain-equation carriers are potentially susceptible to a "hard function change" in the sense of exhibiting a radically new mode of functioning that arises suddenly and effectively irreversibly. This mathematical phenomenon predictably occurs between two brain-equation carriers if they are functionally coupled in a certain simple fashion. Specifically, it occurs when in two brain-equation carriers that are in visual contact, two different social displays—that of the bonding subpotential ("affection") and that of the sum-potential ("happiness")—are sufficiently similar to both act as a bonding reward.

This constellation can be arranged-for artificially between artificial or natural brain-equation carriers, but it can also occur spontaneously through an evolutionary accident called "Huxley ritualization" [13]. The only further precondition is that the two added universal simulators (VR-machines) are sufficiently powerful to allow for "mirror competence" (as the capability in question is called in biology following Wolfgang Kohler and Gordon Gallup). Then, the two autonomous optimizers with cognition will predictably start "cross-optimizing." The positive feedback which occurs in this case does not end in an ordinary "bonding bout" (as in the absence of mirror competence) but spreads to all sources of potential reward for both in a creative fashion.

The consequence is an "interactional function change": Either autonomous optimizer is confronted with the phenomenon of "benevolence" encountered and, in response, begins to attempt to act benevolently, too. This at first only budding but then suddenly taking-over simulational mode (functional attractor) leads to an autocatalytic bonding bout of maximum strength on either side that comprises all momentarily possible acts and objects as tokens. This epigenetic transformation is the "personogenetic function change" (cf. [13, 15]). It implies the sudden competence to invent and/or muster a language and to speak in a universe of speaking persons. This renders the predicted dynamical phenomenon—the "person attractor" as Detlev Linke called it—of great interest to the humanities. Niklas Luhmann wrote on this [16]. Noam Chomsky participated in a long telephone conversation on this

question in 1975. Gregory Bateson accepted the function-change result right after its first publication in the San Diego Biomedical Symposium [17]. Of course, it must be kept in mind that computer technology was far from being up to the task of implementing the brain equation along with its universal simulator at the time. To date, this level ought to be available, as predicted in [5].

Let me also point to the therapeutic potential of the two-coupled-brain-equations paradigm. A straightforward therapy for early childhood autism is implicit in the mathematical picture of two coupled brain equations [7]. The therapy consists in the bonding caretaker adopting a simple—if not easy to maintain—rule: to deliberately produce an acoustic (say) bonding signal whenever momentarily happy in the interaction. This proposed therapy is not confined to the playroom. It is so general it can be extended towards other—terrestrial or non-terrestrial, biological or non-biological—bonding mirror-competent brain-equation carriers. Steven Spielberg's "A.I." of 2001 represents a fictional anticipation.

A second applied implication of the brain equation is not therapeutic but preventive in kind. It enables a rational understanding of René Spitz's and Mary Ainsworth's independent 1945 discoveries of the often lethal disease in young children called "hospitalism." The latter occurs when they are deprived of a bonding partner. Knowledge about the epigenetic personogenetic function change and its transformative power thus can help the weakest of all creatures—illiterate infants—avoid despair where it is the most vital to be spared. Young children in this way suddenly become objects of scientific attention in a mathematics-borne context. The function-change model implies that they are even bigger in their mental stature than adults. The brain equation enables this insight. It elevates the previously ill-defined phenomenon of personogenesis to the status of a whole new field of science—that is, to being accessible to empirical investigation, prevention and therapy. The first documentary recording of the predicted personogenetic transformation of a young human being will be a tear-jerker on Youtube. On the other hand, the fact that this document of a predicted bifurcation has been lacking up until now can also be understood: it attests to the tactfulness of human mothers and fathers—that they sensed that this "holy of holies" of humanity is not a thing for the public eye. It indeed belongs into the sanctuary of the medical profession. But it at the same time also belongs into the tradition of the enlightenment since everyone is a part of the caring—responsible—side of humanity. Only the rational framework enabling this "jump" in public awareness was lacking so far in the absence of the brain equation and its correct understanding.

May I add one more step in the direction of applications? I am referring to cribs. We live in an age in which women have been discovered to be perfectly able to fulfill all male roles—the "second American pioneer age" is a global achievement. What is still lacking in this context is the insight into the deficient competence of the male part of humanity in those fields in which the female part is endowed with a superhuman strength—such as needed already in child birth. While some males also have a strong tendency to "wipe buttoms" as Lorenz said, the power of this almost masochistic trait, present in female nature under the onslaught of the bonding hormone oxytocin (which also controls labor) is totally underrated in public consciousness. "Mothers are the better mothers" for feeling less burdened by a heavy workload due to a stronger

bonding reward obtained by them. Young women are not being informed about this fact any more. Think of the incredibly uncomfortable bodily positions adopted for incredibly long periods on a stretch by nursing mice and cats (if this comparison can be understood in the right way). This trait of an almost superhuman strength, given free of charge to the second human subspecies (the third being children) is unknown to society at large. Females can replace males almost everywhere but the opposite is not true: the males have their deficiencies. In particular, the gentle females cannot imagine that males almost totally lack this pseudo-masochistic trait, as they would tend to call it.

Cribs—even day-care cribs—are not just substitutes for maternal and paternal supervision and handling as an industrialized way to rationalize childcare (which is legitimate in principle). They also have the potential to interfere with an "event" that society does not yet know exists: the Personogenetic Function Change, a mathematical phenomenon belonging into bifurcation theory. Without the brain equation, this unique example of a dynamical function change could not have been spotted by the eye of science. Cribs can harm young children in the core of their prospective personalities by interfering with the P.F.C. Following in the footsteps of Friedrich Frobel, inventor of the "Kindergarten" in the early 19th century, a scientific theory of motherhood and parenthood and brotherhood can be formulated based on the mathematical theory of the brain equation.

7 Conclusion

The brain equation is a surprise finding in mathematics where it occupies a fairly high ("pure") status. Nevertheless it enables a return by society to a responsible dealing with children. At the same time it enables (via the mathematical notion of optimality theory in the sense of the Rashevsky-Rosen school [17]) a rational understanding of benevolence and of what it means to be a person. In this way, the notion of benevolence ("intentional cross-optimization") enters science. It also—if I may add this here since it is so rare that so delicate a topic comes up—offers an occasion in the course of modern enlightenment to appreciate or even see the benevolent character of the two manifest miracles in physics, color and the Now. Of these two, only the latter can be reproduced in artificial implementations of the brain equation. Both "assignment conditions" (added to Newton's "laws" and "initial conditions" [18]) represent provable gifts from without. Mathematics thereby re-gains its place in the humanities. After beauty got re-claimed on the basis of mathematics by my friend Benoit Mandelbrot, now also the good appears to be transportable back into society by mathematics.

Acknowledgments I thank Ali Sanayei, Ivan Zelinka, Kunihiko Kaneko, Klaus Giel, Niels Birbaumer, Friedrich Kümmel, Thimo Böhl and the late Yrish Sefla for discussions. For J.O.R.

Appendix A

The brain equation, taken from [5]:

$$s(\beta) = \alpha \sum_i f^+(p_{iclos}) - \sum_i f^-(p_{iclos}) \tag{A1}$$

with

$$f^+ = g(\phi_i) t_i \, (1 - t_i) \, K \, (x_i + K)^{-1} \tag{A2}$$

$$f^- = \left(1 - \cos \phi_i\right) t_i^2 \left\{ \left(\frac{1 - t_i}{1 - t_i - x_i} \right)^{\gamma} - 1 \right\} \tag{A3}$$

$$\alpha = \left(a + b \sum_i f^- \right)^{-1} \tag{A4}$$

Here, s is the direction-specific weight (sum force), β is a polar angle from zero to 360°, p_{iclos} is the closest source of type i, ϕ_i is the momentary angular difference between β and the direction of p_{iclos}, g is the circle-of-Thales function ($\beta = \cos \phi_i$ if $\phi_i < 90°$ in either direction, and zero outside), the x_i's are the momentary distances between the (constant-speed) traveler and the p_{iclos}, t_i is the remaining time until alarmclock i goes off; and $a = b = 1/4$, $K = 0.001$, $\gamma = 3$ (e.g.).

References

1. Rössler, O.E.: Adequate locomotion strategies for an abstract organism in an abstract environment: a relational approach to brain function. In: Physics and Mathematics of the Nervous System (M. Conrad, W. Guttinger and M. DalCin, eds.), Lecture Notes in Biomathematics, vol. 4, pp. 342–369. Springer, New York (1974)
2. Rössler, O.E., Andreeva, G.: Conjecture: Gödel = lim NP for n → infinity. In: Issues in the proof that P≠NP. http://rjlipton.wordpress.com/2010/08/09/issues-in-the-proof-that-p%E2%89%A0np (2010). Accessed 13 Oct 2010
3. Garey, M.R., Johnson, D.S.: Computers and Intractability: a Guide to the Theory of NP-Completeness. Freeman, New York (1979). http://www.di.unipi.it/~luccio/GJCap1.PDF
4. Rössler, O.E.: Deductive biology–some cautious steps. Bull. Math. Biol. **40**, 45–58 (1978). First 2 pages. http://link.springer.com/article/10.1007%2FBF02463129#page-1
5. Rössler, O.E.: An artificial cognitive-plus-motivational system. In: Rosen, R. (ed.) Progress in Theoretical Biology, Academic Press, New York, vol. 6, pp. 147–160 (1981)
6. Rössler, O.E.: Recursive evolution. BioSystems **11**, 193–199 (1979). Abstract on. http://www.ncbi.nlm.nih.gov/pubmed/497369
7. Rössler, O.E.: Is benevolence compatible with intelligence? On the theory of the humane affection (in German). In: Der Themenppark der Expo 2000, vol. 1, pp. 157–163. Springer-Verlag Vienna (2000)

8. Charnov, E.L.: Optimal Foraging: some theoretical explanations. University of Washington 1973; idem: Optimal foraging: the marginal value theorem. Theor. Popul. Biol. **9**, 129–136 (1976).http://academic.brooklyn.cuny.edu/biology/jbasil/documents/MarginalValueCharnovOptimality.pdf
9. Lorenz, K.: Behind the Mirror: a Search for a Natural History of Human Knowledge. Harcourt, Batavia (1977)
10. Rössler, O.E.: An artificial cognitive map system. BioSystems **13**, 203–209 (1981). Abstract on: http://www.sciencedirect.com/science/article/pii/0303264781900617
11. Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U., Fehr, E.: Oxytocin increases trust in humans. Nature **435**, 673–676 (2005). http://www-psych.stanford.edu/knutson/ans/kosfeld05.pdf
12. Seaman, B., Rössler, O.E.: Neosentience: The Benevolence Engine. In: Intellect Press, Chicago (2011). http://books.google.de/books?id=wfOBzvKCc5IC&printsec=frontcover&hl=de&q&f=false
13. Rössler, O.E.: Nonlinear dynamics, artificial cognition and galactic export (Chaos Award). In: Dubois, D.M. (ed.) Proceedings of Conference on Computing Anticipatory Systems, AIP Conference, vol. 718, pp. 47–67. Melville (2004). http://www.lampsacus.com/documents/roesslergalacticexport.pdf
14. Rosen, R.: Anticipatory Systems - Philosophical, Mathematical, and Methodological Foundations. Pergamon Press, New York (1985)
15. Rössler, O.E.: Interactional bifurcations in human interaction - a formal approach. In: Tschacher, W., Tschiepek, G., Brunner, E.J. (eds.) Self-organization and Clinical Psychology, pp. 229–236. Springer, Heidelberg (1992). First 2 pages. http://link.springer.com/chapter/10.1007%2F978-3-642-77534-5_12#page-1
16. Luhmann, N.: Social Systems: Blueprint of a General Theory (in German), p. 170. Frankfurt am Main: Suhrkamp, Frankfurt (1984). http://steffenroth.files.wordpress.com/2012/03/soziale-systeme.pdf
17. Rössler, O.E.: Mathematical model of a proposed treatment of early infantile autism: facilitation of the "dialogical catastrophe" in motivation interaction. In: San Diego Biomedical Symposium vol. 14, pp. 105–110 (1975)
18. Rössler, O.E., Rössler, J.O.: The endo approach. Appl. Math. Comput. **56**, 281–287 (1993)

Nature Versus Nurture in Complex and Not-So-Complex Systems

D. L. Stein and C. M. Newman

Abstract Understanding the dynamical behavior of many-particle systems both in and out of equilibrium is a central issue in both statistical mechanics and complex systems theory. One question involves "nature versus nurture": given a system with a random initial state evolving through a well-defined stochastic dynamics, how much of the information contained in the state at future times depends on the initial condition ("nature") and how much on the dynamical realization ("nurture")? We discuss this question and present both old and new results for low-dimensional Ising spin systems.

Keywords Heritability · Persistence · Aging · Damage spreading · Ising spin dynamics

1 Introduction

The nonequilibrium dynamics of both thermodynamic and complex systems (the intersection of these two sets is nonempty) remains an area of intensive research, and a host of open problems remains. The most extreme case of nonequilibrium dynamics occurs after a *deep quench*, in which a system in equilibrium at a very high temperature is instantaneously cooled to a very low temperature, after which it evolves according to a well-defined dynamics corresponding to that low temperature.

D. L. Stein (✉) · C. M. Newman
Courant Institute of Mathematical Sciences, New York University,
New York, NY 10011, USA
e-mail: daniel.stein@nyu.edu

D. L. Stein
Department of Physics, New York University, New York, NY 10003, USA

C. M. Newman
Department of Mathematics, University of California, Irvine, CA 92697, USA

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_6,

The extreme case of a deep quench is the instantaneous cooling of a system from infinite to zero temperature. The subsequent zero-temperature dynamics consists of the system's running "downhill" in energy (or uphill in survival probability, if one is dealing with a biological system) to some local or global minimum (or maximum).

Determining the state of such a system at long times, given both the initial state and the subsequent dynamics, is a difficult—and generally unsolved—problem, even for relatively simple systems. In this paper, we will review progress on this question for Ising spin systems, both homogeneous and disordered, in one and two dimensions. We will see that even in $2D$ the problem is far from simple, with open questions remaining even for—in fact, especially for—the uniform ferromagnet. However, recent progress has been made, and the insights gained may be useful in understanding dynamical properties of more interesting—and possibly complex—systems.

2 Types of Long-Time Behavior

For concreteness we consider an Ising spin system on the infinite lattice $\mathbf{Z}^d$; that is, at each site $x \in \mathbf{Z}^d$ we assign a binary variable $\sigma_x = \pm 1$. We restrict our attention to models in which the spin-spin couplings are nearest-neighbor. The most basic question one might ask is whether, after a deep quench, the dynamics eventually settles down to a fixed state, or whether some or all spins continue to flip forever.

The notion of equilibration of an infinite system after a finite time contains some subtleties, which we will address in the next section. But without addressing these subtleties, we can pose the question in a precise way: does the spin configuration have a limit as $t \to \infty$? Equivalently, for every x, does σ_x flip infinitely often or only finitely many times? (Note that even for those systems in which the latter is true, it will generally *not* be the case that there exists some finite time T_0 after which every spin has stopped flipping. This is discussed further in the next section.) From this perspective, it is useful to distinguish among three classes of dynamical system: a system is type $\mathcal{F}$ if every spin flips only finitely many times; type $\mathcal{I}$ if every spin flips infinitely often; and type $\mathcal{M}$ (for "mixed") if some spins flip infinitely often and others do not [1]. The overall spin configuration has a limit only of the system is type $\mathcal{F}$.

Determining which class a system belongs to is generally a nontrivial problem; in fact, the answer remains unknown even for uniform ferromagnets (and antiferromagnets) in $\mathbf{Z}^d$ for $d \geq 3$. There does exist some numerical work, however, suggesting that these might be type $\mathcal{I}$ for $d = 3$ and 4, but type $\mathcal{F}$—or possibly $\mathcal{M}$—for $d \geq 5$ [2].

However, some progress has been made. There exist proofs that uniform ferromagnets or antiferromagnets in one dimension (on $\mathbf{Z}$) and in two dimensions (on $\mathbf{Z}^2$) are type $\mathcal{I}$. Moreover, in *any* dimension on a lattice where each site has an odd number of nearest neighbors, they are type $\mathcal{F}$ [3]. Work has also been done on two-dimensional "slabs": that is, systems that are infinite in two dimensions but consist of a finite number of layers in the third. Here the system can be either type $\mathcal{F}$ or $\mathcal{M}$,

depending both on the number of layers and on the boundary conditions (free or periodic) in the third (finite) direction. For details, see [4].

It was further proved in [3] that all models with continuous disorder, in which the spin-spin couplings are chosen from a common distribution with finite mean, belong to class $\mathcal{F}$ in all dimensions and on all types of lattice. These include ordinary Edwards-Anderson spin glasses [5] and random ferromagnets. We ignore here systems with continuous disorder in which the distribution has infinite mean, and refer the interested reader to [1, 6].

Another class of systems comprises the so-called $\pm J$ spin glass models, where each coupling independently takes on the value $+J$ or $-J$ with equal probability. It was shown that in one dimension these are type $\mathcal{I}$, and in two dimensions (again on $\mathbf{Z}^2$) type $\mathcal{M}$ [7]. And once again, on any lattice regardless of dimension where each site has an odd number of neighbors, they are type $\mathcal{F}$.

Results exist also for systems with more exotic coupling distributions; we refer the interested reader to [1]. We now turn to the next question, which is our main interest here: what can be learned about the state of a system at a finite time t after a deep quench. The answer, not surprisingly, depends on which class the system belongs to, but as we shall see, in most cases one is forced to undertake numerical simulations to gain insight.

3 Local Equilibration, Local Non-Equilibration, and Chaotic Size Dependence

How might one think about equilibration in an infinite system, even one of type-$\mathcal{F}$, given that at any finite time some spins still not have reached their final state? It was proposed in [8] that this problem could be understood in the sense of *local equilibration*: choose a region of fixed size surrounding the origin, and ask whether, after a finite time, domain walls cease to sweep across the region, overturning the spins within. This timescale $\tau(L)$ is expected to increase without bound as L goes to infinity (and in general will also depend on the choice of initial condition, dynamics, lattice type and dimensionality, and possibly other factors); but the idea is that as long as $\tau(L) < \infty$ for *any* $L < \infty$, no matter how large, then we can say that the system undergoes local equilibration. Any system of type-$\mathcal{F}$ obviously undergoes local equilibration. Types $\mathcal{I}$ and $\mathcal{M}$ do not, and we say that these systems experience *local nonequilibration* (LNE) [8].

LNE can be of two types. Even though the configuration in a given finite region never settles down, one can still ask whether, if one averages over all dynamical realizations, the *dynamically averaged* configuration settle down to a limit at large times. Or does even this averaged configuration not settle down?

The first possibility (a limit of the dynamically averaged configuration) can be thought of as "weak LNE", while the second (no limit) is referred to as *chaotic time dependence* (CTD) [8]. As shown in [8], weak LNE implies a complete lack of

predictability (nurture "wins"—after some time, the dynamics wipes out information about the initial state), while CTD implies that some amount (which can be quantified) of predictability remains (nature wins).

So a study of nature versus nurture provides a great deal of information on a number of central dynamical issues concerning classes of dynamical systems. We now review both older and more recent results for different Ising-like spin systems, both homogeneous and disordered.

4 Nature Versus Nurture in $1D$ Random Ferromagnets and Spin Glasses

Because type-$\mathcal{F}$ models always equilibrate locally, one can simply compare the final state of a spin with its initial state over many dynamical trials to determine whether initial information has been fully retained, partially retained, or completely lost. This can be quantified by introducing [3] a type of dynamical order parameter, denoted q_D, that in some ways serves an analogue of the (equilibrium) Edwards-Anderson order parameter q_{EA} [5].

Let σ^t denote the (infinite-volume) spin configuration at time t given a specific initial configuration σ^0 and dynamical realization ω (for notational convenience, the dependence of σ^t on σ^0 and ω is suppressed). We want to study, for fixed σ^0 (and, if the model is disordered, fixed coupling realization $\mathcal{J}$), this quantity averaged over all dynamical realizations up to time t; denote such an average by by $\langle\cdot\rangle_t$. One then needs to study the resulting quantity averaged over all initial configurations and coupling realizations. Denoting the latter averages (with respect to the joint distribution $P_{\mathcal{J},\sigma^0} = P_{\mathcal{J}} \times P_{\sigma^0}$) by $\mathbf{E}_{\mathcal{J},\sigma^0}$, we define $q_D = \lim_{t\to\infty} q^t$ (providing the limit exists), where

$$q^t = \lim_{L\to\infty} |\Lambda_L|^{-1} \sum_{x\in\Lambda_L} (\langle\sigma_x\rangle_t)^2 = \mathbf{E}_{\mathcal{J},\sigma^0}(\langle\sigma_x\rangle_t^2) \tag{1}$$

and Λ_L is a d-dimensional cube of side L centered at the origin. The equivalence of the two formulas for q^t follows from translation-ergodicity [3].

The order parameter q_D measures the extent to which σ^∞ is determined by σ^0 rather than by ω. It was proved in [3] that for the $1D$ random ferromagnet and/or spin glass with continuous disorder, $q_D = 1/2$. What this means is that, for a.e. $\mathcal{J}$ and σ^0, precisely half of the x's in $\mathbf{Z}$ have σ_x^∞ completely determined by σ^0 with the other σ_x^∞'s completely undetermined by σ^0.

We turn now to the more difficult case of type-$\mathcal{I}$ systems. It seems somewhat counterintuitive that models with continuous disorder, in particular random ferromagnets and spin glasses, whose equilibrium thermodynamics are much more difficult to ascertain than those of uniform ferromagnets, are (at least in some cases) *easier* to analyze in the context of nature versus nurture.

5 Persistence and Heritability in Low-Dimensional Uniform Ferromagnets

The nature versus nurture question is intimately related to older notions of *persistence* [9], defined as the fraction of spins that are unchanged from their initial values at time t. This was found to decay as a power law in a number of systems, in particular uniform ferromagnets and Potts models in low dimensions, and the associated decay exponent θ_p is known as the "persistence exponent".

In a similar manner, one can define a "heritability exponent" [10] as follows: prepare two Ising systems with the same initial configuration but then allow them to evolve independently using zero-temperature Glauber dynamics. The spin overlap between these "twin" copies, with the same initial condition but two different dynamical realizations, was found (after averaging over many trials and different initial conditions) to decay as a power law in time [10]. This spin overlap, which we refer to as the "heritability", is essentially the same as q^t. The exponent θ_h associated with the power-law decay of heritability is the "heritability exponent".

Heritability defined in this way is in some sense the opposite of "damage spreading" [11–13]; the latter involves starting with two slightly different initial configurations and letting them evolve with the *same* dynamical realization. The extent of the spread of the initial difference throughout the system is then measured.

The persistence and heritability exponents can be computed exactly in the $1D$ uniform Ising ferromagnet. It was shown in [14, 15] that $\theta_p = 3/8$ for this system. On the other hand, it can be shown that $\theta_h = 1/2$, as discussed in [10], by using the mapping to the voter model and coalescing random walks (see, e.g., [15, 16]).

While the persistence and heritability exponents differ in one dimension, they may be identical in the $2D$ uniform ferromagnet, where numerical simulations yield $\theta_p = 0.21 \pm 0.02$ [2, 17] and $\theta_h = 0.22 \pm 0.02$ [10]. Whether the two exponents are exactly the same, or simply close but not identical, remains to be understood.

6 Positive Temperature

Does the preceding discussion have anything to say about what happens at nonzero temperature? Here one needs to study the behavior of positive temperature Gibbs states and the local order parameter, rather than that of single spin configurations. Construction of the appropriate dynamical measures, analysis of their evolution, and relation to pure state structure are extensively discussed in [8]. Here we mention only a few relevant results.

The categorization into types $\mathcal{I}$, $\mathcal{F}$, and $\mathcal{M}$ is specifically tailored to zero temperature and needs to be modified at positive temperature. In the latter case, one can still define local equilibration, in the sense that, on any finite lengthscale, the system equilibrates into a *pure state* after a finite time (depending on all of the usual culprits), in the sense that interfaces cease to move across the region after that time. If finite regions exist without a corresponding finite equilibration timescale, then LNE occurs.

A main result of [8] is relevant to spin glasses in particular: if only a single pair, or countably many pairs (including a countable infinity) of pure states exists (with fixed $\mathcal{J}$), and these all have nonzero Edwards-Anderson (EA) order parameter [5], then LNE occurs. A corollary is that if LNE does *not* occur, and the limiting pure states have *nonzero* EA order parameter, then there must exist an *uncountable* infinity of pure states, with almost every pair having overlap zero.

One consequence of these results is that LNE occurs at positive temperature (with $T < T_c$) in the $2D$ uniform ferromagnet and (presumably) random Ising ferromagnets for $d < 5$. Because the number and structure of pure states at positive temperature in Ising spin glasses is unknown for $d \geq 3$ (and, from a rigorous point of view, unproved even for $d = 2$), occurrence of LNE there remains an open question.

7 Open Problems

The behavior of homogeneous and disordered Ising spin systems in one and two dimensions is now relatively well understood. Beyond that, however, most questions remain open. Do uniform ferromagnets belong to class $\mathcal{F}$, $\mathcal{I}$, or $\mathcal{M}$ in dimension three and higher? If $\mathcal{F}$, what is the value of q_D? If not, is weak LNE or CTD displayed, and what is the value of the heritability exponent?

The relationships among heritability, persistence, and damage spreading form an interesting set of open problems as well. Are the heritability and persistence exponents the same in the $2D$ ferromagnet on a square lattice, and if so, why? What about higher dimensions and other models? It would be interesting to study these relations in two and higher dimensions and work out the connections between these different but related quantities.

Acknowledgments The authors thank Jon Machta, Jing Ye, Vladas Sidoravicius, and P. M. C. de Oliveira for fruitful collaborations on questions of nature versus nurture. This research was supported in part by US NSF Grants DMS-1207678 and OISE-0730136.

References

1. Newman, C.M., Stein, D.L.: Zero-temperature dynamics of ising spin systems following a deep quench: results and open problems. Physica A **279**, 159–168 (2000)
2. Stauffer, D.: Ising spinodal decomposition at $T = 0$ in one to five dimensions. J. Phys. A **27**, 5029–5032 (1994)
3. Nanda, S., Newman, C.M., Stein, D.L.: Dynamics of ising spin systems at zero temperature. In: Minlos, R., Shlosman, S., Suhov, Y. (eds.) On Dobrushin's Way (from Probability Theory to Statistical Physics). Amer. Math. Soc. Transl. vol. 198 pp. 183–194 (2000)
4. Damron, M., Kogan, H., Newman, C.M., Sidoravicius, V.: Coarsening in $2D$ slabs (2013). http://arxiv.org/abs/1303.2505 (2013)
5. Edwards, S., Anderson, P.W.: Theory of spin glasses. J. Phys. F **5**, 965–974 (1975)

6. Nanda, S., Newman, C.M.: Random nearest neighbor and influence graphs on Z^d. Random Struct. Alg. **15**, 262–278 (1999)
7. Gandolfi, A., Newman, C.M., Stein, D.L.: Zero-temperature dynamics of $\pm J$ spin glasses and related models. Commun. Math. Phys. **214**, 373–387 (2000)
8. Newman, C.M., Stein, D.L.: Equilibrium pure states and nonequilibrium chaos. J. Stat. Phys. **94**, 709–722 (1999)
9. Derrida, B., Bray, A.J., Godreche, C.: Nontrivial exponents in the zero temperature dynamics of the $1D$ ising and potts models. J. Phys. A Math. Gen. **27**, L357–L361 (1994)
10. Ye, J., Machta, J., Newman, C.M., Stein, D.L.: Nature versus nurture: predictability in zero-temperature ising dynamics. Phys. Rev. E **88**, 040101 (2013)
11. Creutz, M.: Deterministic ising dynamics. Ann. Phys. **167**, 62–72 (1986)
12. Stanley, H.E., Stauufer, D., Kertesz, J., Hermann, H.: Dynamics of spreading phenomena in two-dimensional ising models. Phys. Rev. Lett. **59**, 2326–2328 (1987)
13. Grassberger, P.: Damage spreading and critical exponents for model A ising dynamics. Physica A **214**, 547–559 (1995)
14. Derrida, B., Hakim, V., Pasquier, V.: Exact first-passage exponents of $1D$ domain growth: relation to a reaction-diffusion model. Phys. Rev. Lett. **75**, 751–754 (1995)
15. Derrida, B., Hakim, V., Pasquier, V.: Exact exponent for the number of persistent spins in the zero-temperature dynamics of the one-dimensional potts model. J. Stat. Phys. **85**, 763–797 (1996)
16. Fontes, L.R., Isopi, M., Newman, C.M., Stein, D.L.: Aging in $1D$ discrete spin models and equivalent systems. Phys. Rev. Lett. **87**, 110201-1–110201-4 (2001)
17. Jain, S.: Zero-temperature dynamics of the weakly disordered ising model. Phys. Rev. E **59**, R2493–R2495 (1999)

Complex Self-Reproducing Systems

Roderick Edwards and Aude Maignan

Abstract Cellular automata and L-Systems are well-known formal models to describe the behaviour of biological processes. They are discrete dynamical systems, each of which can have complex and varied behaviour. Here, we study a class of substitutive systems incorporating properties of both cellular automata and L-systems, that exhibits self-reproducing behaviour. A one-dimensional array of cells is considered, each cell has a set of modes or states which are determined by a number from $\mathbb{Z}/\mathbf{n}\mathbb{Z}^*$ (n prime). The behaviour of a cell depends on the states of its neighbours and obeys an additive rule. It has also a cell-division mode, which allows the line of cells to grow. The behaviour of such a model can be complex, but, using algebraic techniques, we prove that it can describe a reproducing system.

Keywords Self-reproducing systems · Self-organizing systems · Cellular automata · Substitution systems · L-systems

1 Introduction

Biological processes [3, 7] are often modelled by a continuous formalism in which the behaviour of the system is described by differential equations. The smooth variation of each variable is described as a function of others [4, 9, 14]. Another type of modelling uses a discrete formalism called multi-agent formalism. Cellular Automata (CA) and L-Systems are parts of multi-agent systems that are homogeneous (all

R. Edwards
Department of Mathematics and Statistics, University of Victoria,
PO Box 3060 STN CSC, Victoria, BC, Canada

A. Maignan (✉)
Laboratoire Jean Kuntzmann, 51 rue des Mathematiques, BP 53,
38041Grenoble Cedex 9, France
e-mail: aude.maignan@imag.fr

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_7,

cells are identical). CA are defined by a lattice of sites, an alphabet of symbols and an evolution rule. Each cell evolves in discrete time steps according to some deterministic rules that depend only on local neighbours. In the configuration space, CA trajectories are likely to merge with time, and after many time steps, trajectories starting from almost all initial states become concentrated onto attractors. In these cases, completely disordered starting states evolve to more structured states by a process of self-organisation. In other cases, complex behaviour can appear. In his well-known papers [11, 12], Stephen Wolfram proposed a classification of cellular automaton rules into four classes.

1. Class I: Evolution leads to a homogeneous state.
2. Class II: Evolution leads to a set of separated simple stable or periodic structures.
3. Class III: Evolution leads to a chaotic pattern.
4. Class IV: Evolution leads to complex localised structures, sometimes long-lived.

This classification is based on observations of typical behaviours. Sutner [8], but also Bagnoli [2] and Baetens [1] give a formalisation of those classes. For instance, the elementary rule 30 of Wolfram exhibits many different behaviours of class I, II, III and IV. On the other hand, additive CA with fixed and periodic boundary conditions over a finite field $\mathbb{Z}_{\mathbf{n}}$ are more often of class I and II. This problem of determining dynamical possibilities for additive CA has been studied by Wolfram [10], who has obtained some powerful theoretical results thanks to algebraic tools.

Wolfram has also studied substitution systems [13], discrete dynamical systems such that, at every step, each site is replaced by a fixed block of new sites. Substitution systems are clearly designed so that the number of sites can change. The theory of such systems was first developed by Aristid Lindenmayer and Przemyslaw Prusinkiewicz for modelling plant growth [5, 6], using a framework called L-systems. An L-system is defined as a t-tuple $G = (V, \omega, P)$ where V is an alphabet, ω is the initial state of the system and P is a set of production rules. One of the simplest and best-known examples is the modelling of the anabaena algae. The alphabet is reduced to two states $V = \{A, B\}$ and only two rules are required:

$$\boxed{A \to AB \,|\, B \to A}$$

At each step, each letter A is transformed to the two letter sequence AB and each letter B is transformed to A. If the initial state is A, for example, we obtain the sequence $ABAABABAABAABABAABAABABA$ after 6 steps. The family of possible transformation rules is huge and the study of the behaviours of L-systems is done on a case-by-case basis.

The aim of this chapter is to combine properties of L-systems and additive CA in order to create a simple model which describes a self-reproducing "organism". The algebraic properties of this new model allow one to describe theoretically its behaviour, particularly its reproducing behaviour.

The model is described in Sect. 2. In Sect. 3 we present our analysis of its behaviour.

2 Model Description

We consider a one dimensional structure in which the evolution of a particular site depends on its own value and those of its nearest neighbours. Sites are arranged around a circle (so as to give periodic boundary conditions). This structure is named a **sequence**. The number of sites of a sequence depends on time and is denoted by $N(t)$. The values of the sites of a sequence denoted a also depend on time and are denoted by $a_0^{(t)}, \ldots, a_{N(t)-1}^{(t)}$. The possible site values are elements of a finite set $\mathbb{Z}_{\mathbf{n}}^* = \{\mathbf{1}, \mathbf{2}, ..., \mathbf{n-1}\}$ (n prime). The value of a site of the neighbour-dependent substitution system we consider is computed by an operator $\Phi_{(\alpha,\beta)}$, which depends on two parameters $(\alpha, \beta) \in \mathbb{Z}_{\mathbf{n}}^{\mathbf{2}}$ and is defined by

- a simple additive rule of the form (taking the site index i modulo $N(t)$): $a_i^{(t+1)} = \alpha a_{i-1}^{(t)} + \beta a_i^{(t)} + \alpha a_{i+1}^{(t)}$ [mod n] if $\alpha a_{i-1}^{(t)} + \beta a_i^{(t)} + \alpha a_{i+1}^{(t)}$ [mod n] $\neq 0$,
- and an expansive rule such that $a_i^{(t+1)}$ is divided into two sites of value 1 if $\alpha a_{i-1}^{(t)} + \beta a_i^{(t)} + \alpha a_{i+1}^{(t)}$ [mod n] $= 0$ (i.e., $0 \rightarrow 11$).

We call such a system a **Della Dora system**, because it was first proposed to us by Jean Della Dora. The family of Della Dora systems is parameterised by α, β and n.

Subsequences are open in the sense that, by themselves, they carry no information about what is to either side. Thus we allow the possibility that a subsequence may be a full sequence, with periodic repetitions on either side, though it may, of course, be surrounded by something else.

Because of the expansive rule, the index of a site can change over time. But at a given step, a sequence may be represented by a characteristic polynomial $A^{(t)}(x) = \sum_{i=0}^{N(t)-1} a_i^{(t)} x^i$, and the additive rule can be represented by multiplication of the characteristic polynomial by a fixed Di-polynomial in x, namely $T(x) = \alpha x + \beta + \alpha x^{-1}$, according to $A^{(t+1)}(x) = T(x)A^{(t)}(x) \mod(x^{N(t)} - 1)$. This approach has been developed by Wolfram [10] for a pure additive CA. The new expansive rule modifies (but does not destroy) the algebraic properties of reference [10]. For instance, let us consider the commutative ring $\mathbb{Z}_{\mathbf{3}}$. The sequences built on $\mathbb{Z}_{\mathbf{3}}^*$ are composed of sites of value 1 or 2. Six different dynamics can be defined by the six Di-polynomials $x + x^{-1}$, $2x + 2x^{-1}$, $x + 1 + x^{-1}$, $x + 2 + x^{-1}$, $2x + 1 + 2x^{-1}$, and $2x + 2 + 2x^{-1}$. We present 3 examples which correspond to 3 different behaviours.

Example 1 First, consider the Di-polynomial $T(x) = x + 2 + x^{-1}$ on $\mathbb{Z}_{\mathbf{3}}^*$. The sites evolve such that, at each time step, the value of a site $a_i^{(t+1)}$

- takes the value $a_{i-1}^{(t)} + 2a_i^{(t)} + a_{i+1}^{(t)}$[mod 3] if this value is not equal to 0;
- else, splits into two sites, both of which take the value 1.

This is a specific L-system and can be defined using this following transformation rules, where the site in brackets is modified according to the rule defined by the arrow.

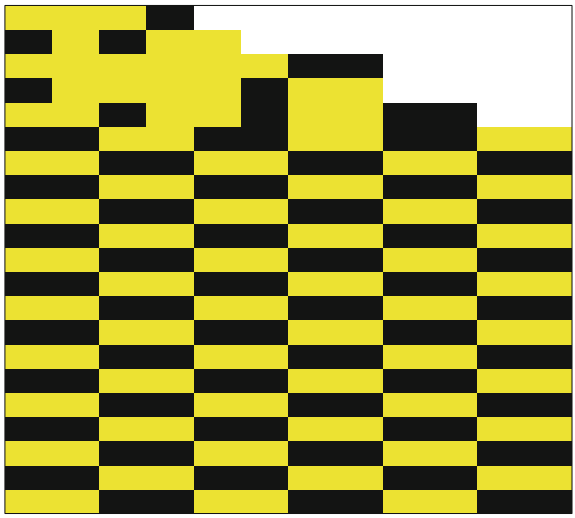


Fig. 1 Example of synchronisation: $T(x) = x + 2 + x^{-1}$ over $\mathbb{Z}_3^*$ with substitution rule $0 \to 11$

$$
\begin{array}{lll}
1\langle 1\rangle 1 \to 1 & 1\langle 1\rangle 2 \to 2 & 1\langle 2\rangle 1 \to 11 \\
1\langle 2\rangle 2 \to 1 & 2\langle 1\rangle 1 \to 2 & 2\langle 1\rangle 2 \to 11 \\
2\langle 2\rangle 1 \to 1 & 2\langle 2\rangle 2 \to 2 &
\end{array}
$$

The evolution of this system can be represented graphically at successive time steps by successive lines. Sites with value one are represented by a light colour; sites with value two are dark.

Figure 1 shows the system's behaviour when the initial configuration is (1112). For every initial configuration, $N(t)$ is bounded and the system converges to a fixed point or a cycle.

This example thus belongs to Wolfram's class *II* and describes a synchronisation phenomenon. This synchronisation is a strong case of a self-organizing process where a global coordination arises out of an initially disordered system.

The behaviour of the second example is more complex.

Example 2 Consider the system with Di-polynomial $T(x) = 2x + 1 + 2x^{-1}$ on $\mathbb{Z}_3^*$ and the $0 \to 11$ rule.

This is defined in detail by the following L-system :

$$
\begin{array}{lll}
1\langle 1\rangle 1 \to 2 & 1\langle 1\rangle 2 \to 1 & 1\langle 2\rangle 1 \to 11 \\
1\langle 2\rangle 2 \to 2 & 2\langle 1\rangle 1 \to 1 & 2\langle 1\rangle 2 \to 11 \\
2\langle 2\rangle 1 \to 2 & 2\langle 2\rangle 2 \to 1 &
\end{array}
$$

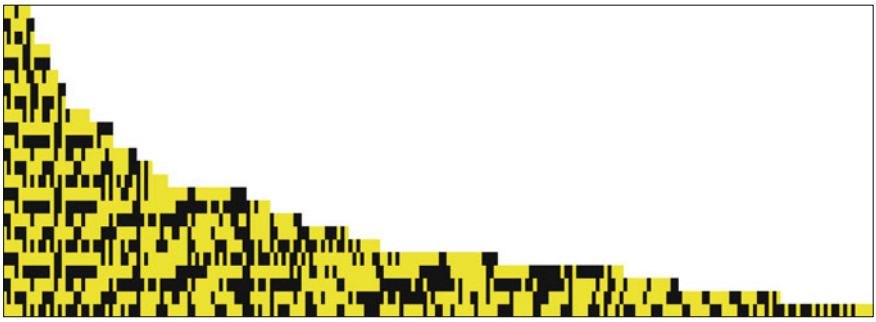


Fig. 2 Example of chaotic behaviour : $T(x) = 2x + 1 + 2x^{-1}$ over $\mathbb{Z}_3^*$, with substitution rule $0 \to 11$

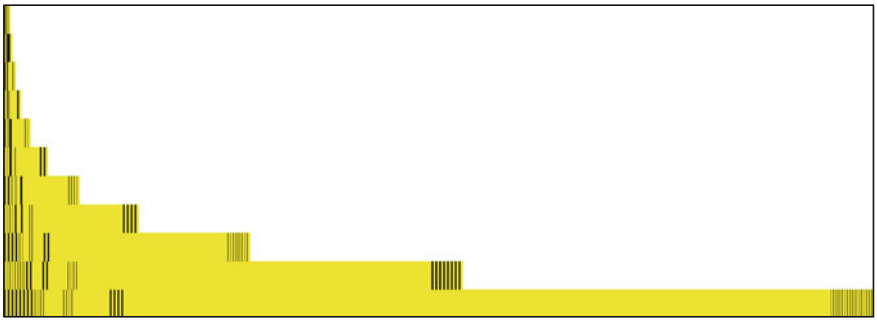


Fig. 3 Example of self replication: $T(x) = 2x + 2 + 2x^{-1}$ over $\mathbb{Z}_3^*$, with substitution rule $0 \to 11$

Figure 2 shows the system's behaviour when the initial configuration is (1112111). The behaviour is complicated and there is no obvious nested structure, so it belongs to Wolfram's class III.

Between the above examples, there exists a class of dynamics where the behaviour is both rich and predictable.

Example 3 Consider the system with Di-polynomial $T(x) = 2x + 2 + 2x^{-1}$on $\mathbb{Z}_3^*$, again with the rule $0 \to 11$.

The transformation rule is described in the following table:

$$
\begin{array}{lll}
1\langle 2\rangle 1 \to 2 & 1\langle 2\rangle 2 \to 1 & 1\langle 1\rangle 1 \to 11 \\
1\langle 1\rangle 2 \to 2 & 2\langle 2\rangle 1 \to 1 & 2\langle 2\rangle 2 \to 11 \\
2\langle 1\rangle 1 \to 2 & 2\langle 1\rangle 2 \to 1 &
\end{array}
$$

The behaviour is complicated but nested structure can be identified. Moreover nested structures give birth to new nested structures. This belongs to Wolfram's class IV. Figure 3 shows the behaviour from inital configuration (1122121211211).

More generally, the discrete phase portrait of such systems is difficult to analyse because the dimension of the phase space evolves over time. Nevertheless, algebraic tools bring powerful results. In Example 3 above, the open subsequences $11...11 = 1^k$ with $k \geq 4$ always produce a subsequence of 1s of at least the same length after each time step (when k is exactly 4, the subsequence of 1s remains the same length: the open subsequence 21^42 produces another sequence containing 21^42). This property is at the root of a very interesting type of behaviour: self-replication. We have, in fact, that

Proposition 1 $a_{i-1}^{(t)} = a_i^{(t)} = a_{i+1}^{(t)} = 1$ *implies that* $a_i^{(t+1)}$ *is transformed into two sites of value* 1 *if and only if* 1 *is a root of its Di-polynomial.*

So in $\mathbb{Z}_{\mathbf{n}}$ ($n > 2$), $\Phi_{(\alpha,\beta)}$ gives behaviours of Wolfram class IV if $(x-1)|x(\alpha x + \beta + \alpha x^{-1})$. Equivalently, $\Phi_{(\alpha,\beta)}$ gives behaviours of Wolfram class IV if $2\alpha + \beta = 0 \pmod{n}$. We call systems with this property, $2\alpha + \beta = 0 \pmod{n}$, **Della Dora systems of type 1**.

Every configuration in a Della Dora systems of type 1 has a unique successor at every time step. However, a configuration may have several distinct predecessors. $\alpha x^2 + \beta x + \alpha$ is not prime and $\Phi_{(\alpha,\beta)}$ is not injective so these dynamics are irreversible.

Section 3 describes the properties of Della Dora systems of type 1. We will see how the behaviours of these systems can be described by a finite number of subsequences. This corresponds to a self-organizing phenomenon. We will also see that over time, specific subsequences give birth to a copy of themselves. This corresponds to a self-reproducing phenomenon.

3 Analysis of Della Dora Systems of Type 1

3.1 First Definitions

Cycles in phase space can be determined by a polynomial approach. A P-periodic cycle of an additive CA is a list of P sequences, each of length N,

$$(a^{(t)}, a^{(t+1)}, ..., a^{(t+P-1)}),$$

such that their corresponding characteristic polynomials, $A^{(t+i)}, (i \in \{0, 1, ...P-1\})$ obey

$$A^{(t+i+1)}(x) = T(x)A^{(t+i)}(x) \mod(x^N - 1)$$

and

$$A^{(t+i)}(x) = T(x)^P A^{(t+i)}(x) \mod(x^N - 1).$$

The cycles of the dynamical system $\Phi_{(\alpha,\beta)}$ are the cycles of the additive dynamics defined by the Di-polynomial $T_{(\alpha,\beta)}$, such that every sequence occurring in the cycle

has no null sites (a 0 site would lead to a splitting $0 \rightarrow 11$ and thus could not be part of a cycle).

For Example 3, it is possible to give a formal proof showing that there are only two types of cycles, one of periodicity $P = 1$ and another of periodicity $P = 2$. The cycles of periodicity 2 can be denoted by $((2211)^k, (1122)^k)$ where k is an integer and $(2211)^k$ and $(1122)^k$ are the successive states of the cycle (i.e., $\Phi_{(2,2)}((2211)^k) = (1122)^k$ and $\Phi_{(2,2)}((1122)^k) = (2211)^k$).

Here, $(2211)^k$ denotes a sequence of length $4k$ such that $(2211)^k = 22112211...2211$. With the same notation, the cycles of periodicity 1 can be denoted by $((21)^k)$. All these cycles are reached only by themselves. In general, it is possible also to have transient approaches to cycles. An example in the system with $\alpha = 2$, $\beta = 1$ and $n = 5$ (a Della Dora system of type 1) is the sequence (141) for which $\Phi_{(2,1)}(141) = (131)$ which is then part of the cycle ((131), (424)).

Over time, an open subsequence $11...11 = 1^k$ ($k \geq 4$) produces a subsequence of 1s with length that grows (exponentially) without bound when $k > 4$, and with length that may remain fixed or may grow when $k = 4$, with no further variation in behaviour. If we consider these as sites in a 'base state,' then information resides in regions in which some 'not-1' sites occur. The **characteristic sequences** of a sequence a are maximal open subsequences of sites of $\mathbb{Z}_{\mathbf{n}}$ which do not contain 1^k with $k \geq 4$ and neither begin nor end with a 1. In effect, the occurrence of expanding subsequences of 1s separate the characteristic sequences, and the subsequence 1111 can be considered a **separator**: if b and c are two characteristic sequences of a sequence a, changing the values of any sites of b will never affect values of sites of c. Characteristic sequences b and c evolve independently. They can have periodic behaviour, they can vanish, and they can also grow. If 1111 appears at some time step from the interior of a characteristic sequence, it creates new independent subsequences. This is, in that case, a reproduction process.

A sequence containing the subsequence 1111 we call a **broken** sequence; otherwise it is **unbroken**. Broken sequences consist of one or more characteristic sequences separated by separators of the form 1^k with $k \geq 4$. Unbroken sequences cannot be considered characteristic sequences, since they do not have 1111 to either side.

3.2 Dynamics of Characteristic Sequences

The dynamics of a characteristic sequence (which is open) is sightly different from the dynamics of a (closed) sequence. Let Φ_o denote this new dynamics applied to an open sequence $s = s_0...s_{N-1}$ surrounded by multiple sites with value 1. Then,

$$\Phi_o(s) = \alpha + \beta + \alpha s_0, \alpha + \beta s_0 + \alpha s_1, \ldots, \alpha s_{i-1} + \beta s_i + \alpha s_{i+1}, \ldots, \alpha s_{N-2} + \beta s_{N-1} + \alpha, \alpha s_{N-1} + \beta + \alpha\,,$$

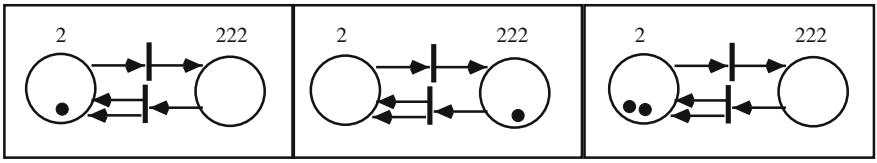


Fig. 4 Petri net of the characteristic sequence (2) at time t, $t+1$ and $t+2$ under the dynamics of the system with Di-polynomial $T = 2x^2 + 2x + 2$ over $\mathbb{Z}_{\mathbf{3}}^*$

where, as before, we apply the expansive rule $0 \rightarrow 11$ where applicable. The new computation of the extremal sites is a consequence of the fact that s is bordered by sites of value 1. The independent evolution of characteristic sequences can now be expressed by $\Phi(b1^*c1^*) = \Phi_o(b)1^*\Phi_o(c)1^*$ where $(b1^*c1^*)$ is a closed sequence constituted by two characteristic sequences, and * is a non specified integer greater than or equal to 4. It is convenient to introduce a more compressed notation: $\Phi_o(b; c) = \Phi_o(b); \Phi_o(c)$. This notation can easily be extended to more than two characteristic sequences.

For Example 3, using this new notation, $\Phi_o^2(2) = \Phi_o(222) = 2; 2$. So, when the characteristic sequence (2) appears, after two steps it splits into two characteristic sequences (2). The periodicity of reproduction of the characteristic sequence (2) is therefore 2.

Figure 4 represents the Petri net of this simple example of behaviour. More complex examples can be represented in this way. For instance, let us consider again the dynamics $\Phi_{(2,1)}$ over $\mathbb{Z}_{\mathbf{5}}^*$. Figure 5 shows the Petri net associated with the characteristic sequence 233213.

The same characteristic sequences can appear several times in a sequence. Let $\mathcal{S}_a$ denotes the set of distinct characteristic sequences that are contained in the sequence a.

For Example 3, it is easy to show that 1^k ($k \in \mathbb{N}$ and $k \geq 4$) are the unique separators. We can consider 1111 a **weak separator**, i.e., a separator that persists but whose length does not increase. Two characteristic sequences that are separated by a weak separator correspond to independent entities which stay linked geographically. For example, $\Phi_{(2,2)}(21^42) = (121^421)$ and the sequence of 1's does not grow longer. On the other hand, if the separator is not weak, the characteristic sequences move apart.

3.3 Behaviour Analysis

For all Della Dora systems of type 1 we have studied, with various choices of α, β and n, we have proven that the behaviour of any sequence that does not belong to a cycle, and is not on a transient leading to a cycle, can be described by a finite number of characteristic sequences.

More precisely, let a be the initial state. Even if the sequence continually grows, there exists a finite set of characteristic sequences $\mathcal{S}$ depending on a, such that for

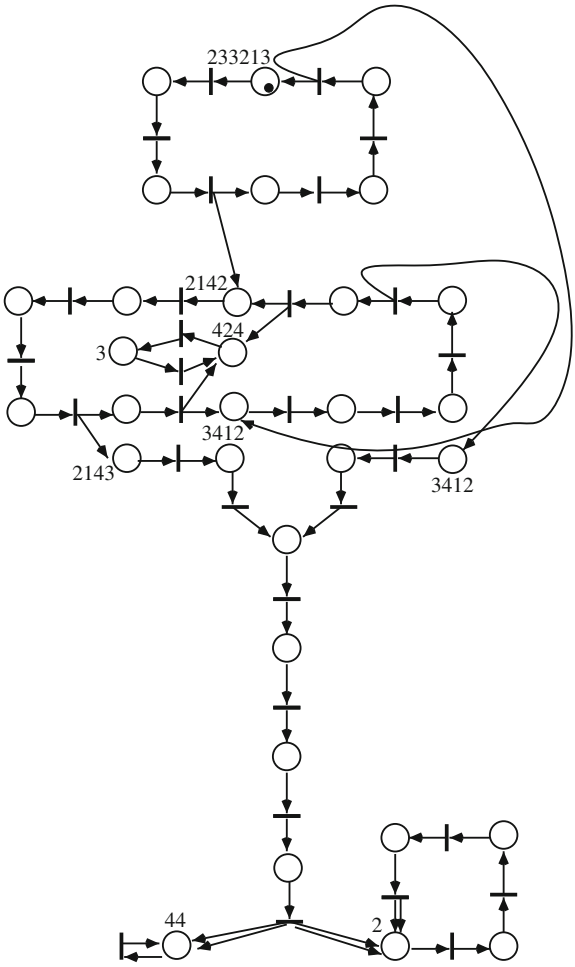


Fig. 5 Petri net for the characteristic sequence (233213) under the dynamics of the system with Di-polynomial $T = 2x^2 + x + 2$ over $\mathbb{Z}_5^*$

all integers l, $\mathcal{S}_{\Phi^l(a)} \subset \mathcal{S}$. In order to prove this strong property, new terms have to be defined.

Broken sequences grow without bound: $\lim_{t\to\infty} N(t) = \infty$. This is true even if there is only one separator and it is weak, because, even though the length of the separating sequence may remain fixed, it feeds new sites into the characteristic sequence at either side. Unbroken sequences may break and then grow without bound, or may not, such as is the case when they fall into a cycle, possibly after a transient. Of course, if a sequence grows without bound, it cannot be part of a cycle or a transient leading to a cycle, and if a sequence does not grow without bound, it must reach a cycle because the accessible state space is then finite. This still leaves open the possibility, *a priori,* of an unbroken sequence that never breaks but grows without bound. In order to deal with both unbroken sequences and characteristic sequences,

and in particular to investigate whether or not they eventually split, we consider the internal dynamics of an arbitrary subsequence.

Let Φ_r denote the partial evaluation of an open sequence s such that:

$$\Phi_r(s) = \alpha s_0 + \beta s_1 + \alpha s_2, \ldots, \alpha s_{i-1} + \beta s_i + \alpha s_{i+1}, \ldots, \alpha s_{N-3} + \beta s_{N-2} + \alpha s_{N-1}$$

where the expansive rule $0 \to 11$ is once again applied. Note that $\Phi_r(s)$ does not transform the end sites of the open sequence s, because that would require information on sites outside of s.

We say that a subsequence s **intrinsically separates** if there exists a number of steps k such that $\Phi_r^k(s)$ contains a separator. When a subsequence separates intrinsically, the value of the first site is never influenced by the value of the site which is just after the subsequence.

Unbroken sequences that eventually cycle clearly do not separate (break). Characteristic sequences can also cycle in the sense that, using our condensed notation, $\Phi_o^P(s) = s$ and these also clearly do not separate. An example, in the system with $\alpha = 2$, $\beta = 1$ and $n = 5$, is $s = 3$ for which $\Phi_o^2(3) = \Phi_o(424) = 3$. Characteristic sequences can also be on transients leading to cycles, such as $s = 4$ in the above example, for which $\Phi_o(4) = 3$. If the characteristic sequence s cycles, there exists an integer i_s ($0 \leq i_s \leq 3$) such that $(s1^{i_s})^k$ also cycles, but $\gamma(s1^{i_s})^k$ ($\gamma \in \mathbb{Z}_{\mathbf{n}}$) does not necessary cycle any more. For instance $(424)^{10}$ cycles but $4(424)^{10}$ does not cycle and does not separate intrinsically. The cyclic part of $\Phi_o^t(4(424)^{10})$ decreases when t increases and $\Phi_o^{10}(4(424)^{10}) = 31124333114; 3; (424)^8$. In other examples, the cyclic part can also totally disappear. We call a **disturbed cycle** a characteristic sequence of the form $*(s1^{i_s})^k*$ where s is a cycle and $*$ represents any subsequence of $\mathbb{Z}_{\mathbf{n}}$ that does not contain a separator, and is small enough such that $*(s1^{i_s})^k*$ does not separate intrinsically. $*$ can also be the empty set, for instance, $s = 11211211324434434423$ is such that $\Phi_{(2,1)}^5(s) = s$ but $\Phi_o^5(s) = 3; 21211122111211324434434311 2$ and s is already a disturbed cycle.

Our objective, then, is to show that every subsequence above a certain size either separates intrinsically or is a transient to a disturbed cycle, or eventually cycles, whether considered as an unbroken sequence or as a characteristic sequence. In the case of intrinsically separable subsequences, even if, over the k steps required to guarantee separation, some of the resulting characteristic sequences became longer, there is still a finite size to which they can grow in k steps. In the case of disturbed cycles, we have observed that, there exists a finite small integer λ such that, for all integers $l > \lambda$, the number of attached cycles s of $\Phi_o^l(*(s1^{i_s})^k*)$ decreases or remains constant and there exists an integer $\hat{l}$ such that $\Phi_o^{\hat{l}}(*(s1^{i_s})^k*)$ is composed of intrinsically separable subsequences and cycles. Thus, over time, a disturbed cycle disappears and generates a set of characteristic sequences. The size of those characteristic sequences is bounded and the set is finite.

This will provide a proof that the number of characteristic sequences generated by any initial sequence is bounded. Consequently, for any given initial broken sequence,

the entire future behaviour can be described by a finite number of characteristic sequences.

We have developed an *R* program whose arguments are n of the finite set $\mathbb{Z}_{\mathbf{n}}^{*}$, α and β of the dynamics and a length N for subsequences. The program computes the list of all the subsequences of length N that do not separate intrinsically using the separator 1111.

For Example 3, running our *R* program shows that every subsequence of length at least 18 either intrinsically separates or is a disturbed cycle or eventually cycles.

Thus, the behaviour of every broken sequence can be defined by a finite set of characteristic sequences.

More generally, we conjecture that for any Della Dora system of type 1, there exists a length N such that every characteristic sequence or unbroken sequence of length N

- intrinsically separates
- or is a transient to a disturbed cycle or a cycle
- or eventually cycles.

This will imply that the entire future behaviour of every sequence can be described by a finite number of unbroken sequences and a finite number of characteristic sequences and the separators. Finally the evolution of any sequence would be representable by a finite Petri net.

4 Conclusion

Cellular automata are capable of very complex behaviour. They model a large range of biological phenomena. For instance, Wolfram has used them to describe natural systems from snowflakes to mollusc shells. The class of Della-Dora systems are less simple in construction than cellular automata and can also have very complex behaviour. The idea is to add the possibility of modelling the important biological process of mitosis. In this extended abstract, we have focussed on the study of a subclass of such systems, which we call Della-Dora systems of type 1. These are irreversible systems that give birth to very interesting self-organizing and self-reproducing behaviours. In future work, the behaviour of Della-Dora system of type 1 will be explored as a simulation of the type of behaviour seen in algae or filaments in a Petri dish. A formalisation for the classification of Della-Dora system of type 1 will also be explored. For Cellular Automata, quantitative measures have been proposed to identify the Wolfram's classes. In our case, the number of sites evolves and these measures must be adapted.

References

1. Baetens J.M., De Baets B.: Phenomenological study of irregular cellular automata based on Lyapunov exponents and Jacobians. Chaos **20**, 033112 (2010)
2. Bagnoli, F., Rechtman, R., Ruffo, S.: Damage spreading and lyapunov exponents in cellular automata. Phys. Lett. A **172**, 3438 (1992)
3. Barnes, D., Chu, D.: Introduction to Modelling for Biosciences. Springer, Berlin (2010)
4. Junck, J.R.: Ten equations that changed biology: mathematics in problem-solving biology curricula. Bioscene 23(1), 11–36 (1997)
5. Lindenmayer, A.: Mathematical models for cellular interaction in development (Parts I and II). J. Theor. Biol. **18**, 280–315 (1968)
6. Prusinkiewicz, P., Lindenmayer, A.: The Algorithmic Beauty of Plants. Springer, New York (1996)
7. Murray, J.D.: Mathematical Biology, 3rd edn, 2 vols. Mathematical Biology: I. An Introduction (2002). Mathematical Biology: II. Spatial Models and Biomedical Applications (2003). Springer, New York
8. Sutner, K.: Classification of cellular automata. In: Meyers, R.A. (ed.) Encyclopedia of Complexity and System Science. Springer, Berlin (2009)
9. Turchin, P.: Evolution in population dynamics. Nature **424**, 257258 (2003)
10. Martin, O., Odlyzko, A.M., Wolfram, S.: Algebraic properties of cellular automata. Commun. Math. Phys. **93**, 219 (1984)
11. Wolfram, S.: Computation theory of cellular automata. Comm. Math. Physics **96**(1), 15–57 (1984)
12. Wolfram, S.: Universality and complexity in cellular automata. Physica **10D**, 1–35 (1984)
13. Wolfram, S.: A new kind of science, Wolfram Media. http://www.ics.uci.edu/eppstein/ca/wolfram.html (2002)
14. Yoshida, T., Hairston, N.G., Ellner, S.P.: Evolutionary tradeoff between defence against grazing and competitive ability in a simple unicellular alga, Chlorella vulgaris. Proc. R. Soc. Lond. B **271**, 1947–1953 (2004)

On Fundamentals of Global Systems Control Science (GSCS)

Raimundas Jasinevicius and Vytautas Petrauskas

Abstract Globalization leads us towards dealing with very complex systems that consist of evolving, overlapping, and interacting "socio-technical fabrics". An existing general systems control theory cannot cope with problems occurring in such systems. This chapter is, first of all, an attempt to present an entirely new approach to the adequacy of system model and reality, based on a causal correspondence between information and knowledge obtained from a reality and its model. Secondly, the chapter suggests two possible control loops: one is meant to improve the model and another is the way to attain a certain planned goal to be reached by our reality. Four doctrines are presented as the basic principles of general fuzzy systems control theory (GFSCT) aiming to deal with the real fuzzy systems operating and functioning in a multiple space-time coordinate system. The minimization of a certain potential V-function is considered as a universal principle for existence of each system in the real world. Moreover, decentralized stochastic control is proposed to improve our reality and guarantee its lifetime unlimited behavior with a proper degree of certainty and space-time stability.

Keywords Globalization · General fuzzy systems control theory · Socio-technical fabric · Adequacy · Data · Information · Knowledge · Wisdom · Uncertainity · Control loop · Space-time coordinate system · Anisotropic space · Decentralized adaptive control · Potential function · Stochastic approximation · Lifetime unlimited stability · Systemology · Methodology · Praxeology

R. Jasinevicius (✉) · V. Petrauskas
Informatics Department, Kaunas University of Technology, Studentu 50-204a,
LT-51368 Kaunas, Lithuania
e-mail: raimundas.jasinevicius@ktu.lt

V. Petrauskas
e-mail: vytautas.petrauskas@ktu.lt

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_8,

1 Globalization: The State-of-the-Art

Our contemporary world in general is too complex and unrecognized yet, and this is the reason why we fight consequences or symptoms instead of dealing with the causes and reasons of global crises, catastrophes or other similar threats [1–9]. We feel the lack of adequate global theories and adequate tools for tackling arising problems.

In reality a lot of feedback and feed forward relations occur, and only such tools as extended fuzzy cognitive maps (FCM) [10], extended fuzzy expert maps (FEM) [11] as well as the approaches of different multi-agent systems (MAS) are able to represent all possible relationships existing among these entities.

Therefore, today's global systems science must answer several key questions:

1Q: How to describe such type of systems using the universal feedback paradigm?
2Q: How to build system models adequate to the reality?
3Q: How to use the results of simulation?

2 Global Systems Model: The Descriptive Approach

First of all, this section tries to emphasize that up to now all more or less serious models of reality have been built using the universal feedback paradigm [12]. In the framework of this paradigm, the Control Theory (CT), Systems Theory (ST), Multi-Agent Systems Theory (MAST), General Fuzzy Systems Control Theory (GFSCT), Collective Adaptive Systems Theory (CAST), Complex Systems Science Theory (CSST) and various other approaches directed to represent formally the most important aspects of reality's behavior emerged as a response to the question 1Q [13, 14].

The global systems model supposedly covering the reality is presented in the Fig. 1.

The set of entities $\Omega = \{\omega 1, \omega 2, \omega 3, \omega 4, \omega 5\}$ has its own objectives and goals.

Those generalized goals can be derived (or in some cases only guessed) from a certain potential function (known or unknown) [15].

Consequently, "the reality" $\Omega = \{\omega 1, \omega 2, \omega 3, \omega 4, \omega 5\}$ under investigation as a whole (and its entities ω) functions according to the second universal paradigm, i.e. the paradigm of causal uncertainty. This paradigm determines an emergent behavior of "the reality" Ω facing all possible risks. According to the ISO, risk is the "effect of uncertainty on objectives" [16].

From this point of view, the main tasks of a global systems model are as follows:

(1) To reduce the uncertainty of long term stability (in "unlimited lifetime") of "the reality" Ω;
(2) To reduce the uncertainty of individual objectives and goals as well as the uncertainty of the emergent Ω;
(3) To improve the quality of reasoning and decision making in Ω under uncertain circumstances;
(4) To reduce the uncertainty of Ω functionality in multiple scale space-time coordinate system.

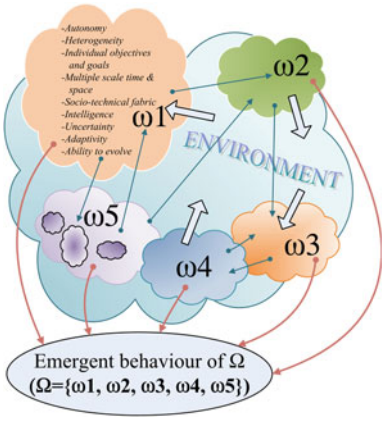


Fig. 1 "Our world"—the reality under investigation

3 Adequacy of the Model and the Reality

In order to answer question 2Q ("How to build system models adequate to the reality?"), everybody looks for an opportunity to compare model data with the corresponding data produced by the reality under investigation. In any case such a comparison covers input (I), output (O) as well as internal state (S) variables wherever they occur: in the entire Ω, in each entity ω, or in their smaller elements. In Fig. 2, the corresponding differences ε (ε_I, ε_S and ε_O) between the reality and its model are determined.

The object of computerized system analysis, decision making processes and information processing shifts from raw data towards more sophisticated computing according to the following scheme [17, 18]:

Data → Information → Knowledge → Wisdom.

As a matter of fact, each step in this transformation scheme is performed on the basis of certain operations. For example, the transformation of raw data (D_R and D_M) into information includes the procedure of data mapping on some context (C_R—in case of a reality, and C_M—in case of a model) [19], as it is shown in Fig. 3.

More sophisticated operations are involved in the process when information is to be transformed into knowledge.

The last step in this scheme (transformation of knowledge into wisdom) is still under discussion and thorough investigation [20, 21].

Under such circumstances, the model adequacy can be evaluated using a special operation $\mathcal{O}$ which performs a comparison of information INF_R and information INF_M and creates a qualitative and/or quantitative measure ε_{INF} (see Fig. 3).

More reliable evaluation of model adequacy to the reality can be obtained on the level of knowledge as shown in Fig. 4.

The reasoning on knowledge level enables us to compare existing theories with the results KN_R received from the investigation of a reality and to correct the theories.

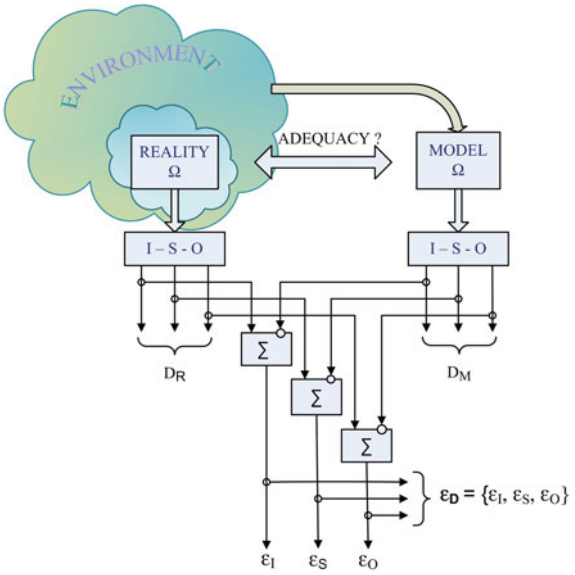


Fig. 2 Evaluation of adequacy of the model and the reality based on data differences

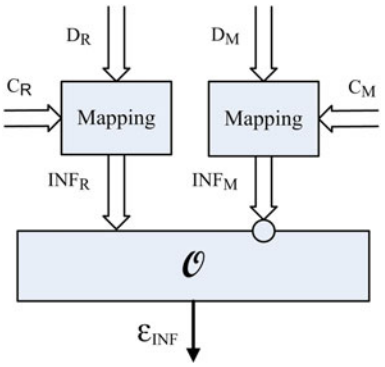


Fig. 3 Raw data transformation into information by mapping on a context

4 Improvement of Model: The First Control Loop

Whatever adequacy measures are obtained (ε, ε_{INF} or ε_{KN}), they all serve as a means to improve the model. It means that in case of weak or uncertain adequacy, the first control loop must be activated with a purpose to correct or adjust functional organization of the model according to the algorithm C_M. Figure 5 demonstrates the case when the difference ε_{INF_1} is used.

The most important theoretical and practical problems to be solved are as follows: a) proposition and research of methodology to determine an operation $\mathcal{O}$ and b) invention and investigation of efficient algorithms C_M.

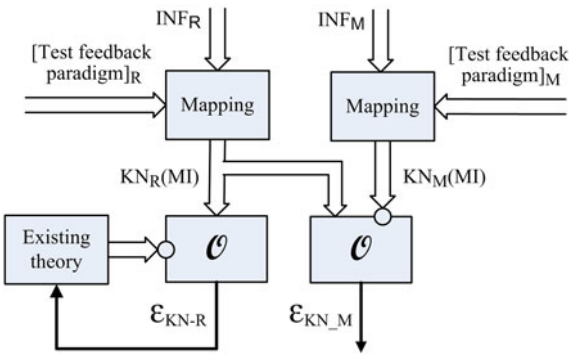


Fig. 4 Adequacy of model and reality evaluation based on knowledge differences

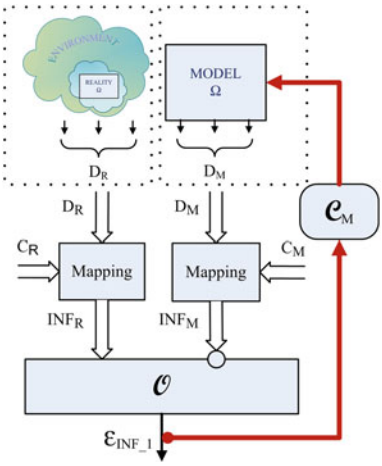


Fig. 5 The first control loop to correct or adjust functional organization of a model

5 Improvement of Reality: The Second Control Loop

The global systems modeling is directed to achieve three main goals: (1) to create a possibility to perform investigation of properties of the reality on its model; (2) to determine a degree of certainty and predictability of its long time stability/instability; and (3) to formulate goals to be achieved by the reality. This must be done according to processes shown in Fig. 6.

As a matter of fact, the most important theoretical and practical problem to be solved on this stage is the invention and investigation of efficient algorithms C_R.

6 Remarks on Global Control

The Systems Theory (ST) permits us to build and to analyze a new model of a real system, using the feedback principle, evaluating and improving its stability, efficiency and other properties only if the real entity to be controlled ω_R and its model ω_M are

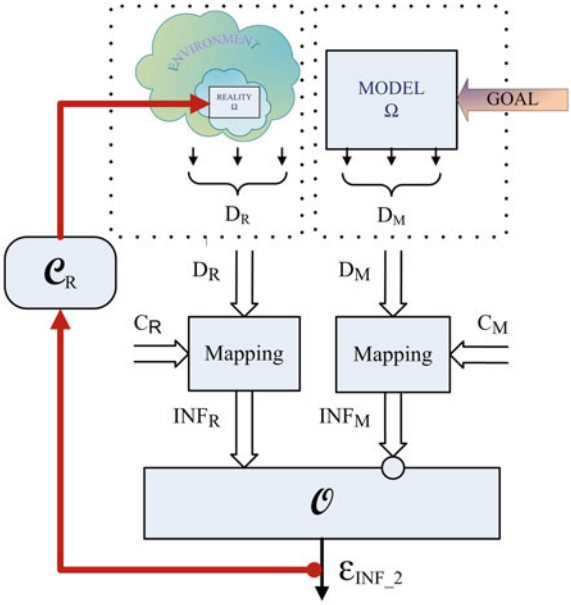


Fig. 6 The second control loop to correct or adjust "the reality", a certain goal to be achieved

relatively very simple. Such generalized examples of the most popular cases are presented in Fig. 7.

Figure 7a corresponds to the functional organization of a typical feedback (FB) control; Fig. 7b—to the feed forward control (FFC) type correction of the output result O_R/O_M and Fig. 7c—to the control functional organization which takes into account not only the difference ε between an entity goal and its result O_R/O_M evaluated through the feedback FB, but also the history of the difference (some sort of its average) and the tendency of changes in ε (some sort of its derivative).

Systems control scientists and experts dealing with systems and entities shown in Fig. 7 are satisfied with the theory of systems stability analysis and with practical tools for time-dependent systems [22].

When "our world" consists of entities characterized by their Autonomy, Heterogeneity, existence of Individual objectives and goals, Activity in multiple scale time and space coordinate system, Intelligence, Uncertainty, Adaptability and its Ability to evolve as well as the possible presence of a human or social factor in those entities the classical Systems theory is powerless [3, 23].

Therefore, a new branch of general systems theory referred to as General Fuzzy Systems Control Theory (GFSCT) is being developed [13]. The GFSCT is based on four main doctrines.

The first doctrine: *In general, our "real world", or our highly interconnected "socio-technical fabric" is intrinsically unstable in the "unlimited lifetime"*

The second doctrine: *In our "real world" and in its models, the decision making and control processes are performed on information and knowledge level using more qualitative than quantitative variables (words and perceptions rather than numbers and measurements)*

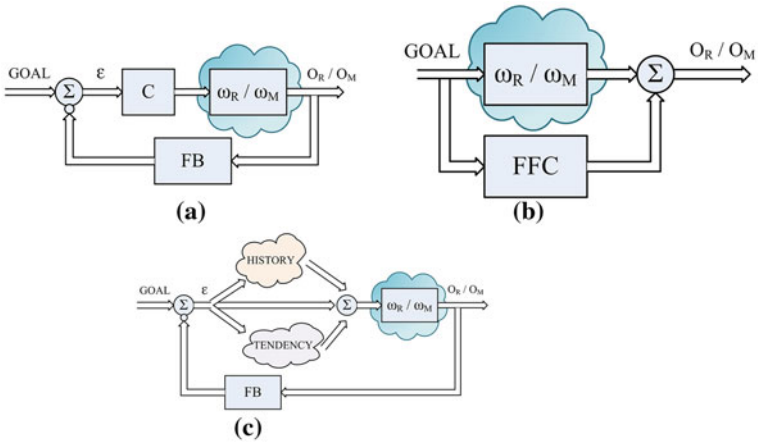


Fig. 7 Generalized examples of a simple feedback (**a**), feed forward (**b**), and complex feedback control (**c**)

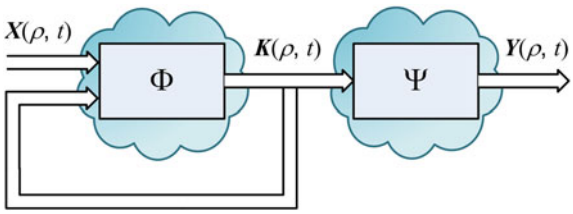


Fig. 8 Transformations of system's internal states and system's outputs

That is the reason why we can witness machine intelligence shifting towards the possibility of human-level performance and computing with words (CWW) [17, 20, 24–26].

The third doctrine: *The "real world" processes and events take part in a multiple scale time and space coordinate system.*

The best attempt to formalize the description of space-time dependent systems was proposed by prof. Wunsch from Dresden University in 1975 [27]. The system Ω was presented by a raw consisting of five sets of variables ***X, K, Y, R, T*** and of two functional transformations $\Phi, \Psi : \Omega = \{\boldsymbol{X, K, Y, R, T}, \Phi, \Psi\}$; here ***X***—a set of inputs I, ***K***—a set of system's internal states S, ***Y***—a set of system's outputs O, ***R***—an independent space variable, and ***T***—an independent time variable; $\Phi: \boldsymbol{X} \times \boldsymbol{K} \times \boldsymbol{R} \times \boldsymbol{T} \rightarrow \boldsymbol{K}$ and $\Psi : \boldsymbol{K} \times \boldsymbol{R} \times \boldsymbol{T} \rightarrow \boldsymbol{Y}$ are transformation of system's internal states and transformation of system's outputs correspondingly (see Fig. 8).

Some significant or decisive processes and events or actions and their influence in time scale have different consequences in case of isotropic and anisotropic spatial relations.

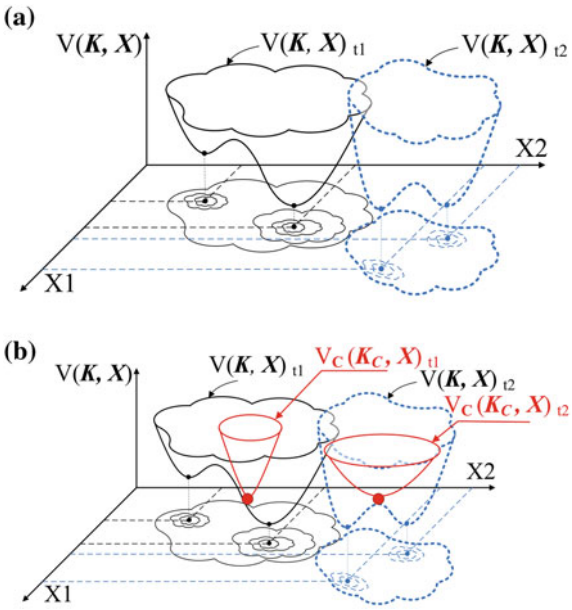


Fig. 9 Target (potential) V-function: a general case (**a**); a stable case (**b**)

The forth doctrine: *To maintain the "real world" predictability of its behavior and a space-time stability/instability must be accompanied by adequate decentralized adaptive control.*

The problem is so complex that it even requires new special measures of coordination and control. Indeed, "Some targeted autocratic industrial policy may yield some seemingly spectacular economic success at some transitional state of development (in a catch-up phase) but at a huge price of externalities and human costs and thus may not be sustainable in the very long run unless they switch to a democratic regime with comprehensive economic and political freedom" (Gottinger in his internet letter to BISC Group members, 16 Dec. 2011).

7 Target (Potential) V-function

Nowadays it is widely agreed that the minimization of a certain target or potential V-function is a universal principle of existence of each system in the real world [15, 22, 28–30]. The latter property is symbolically shown in Fig. 9a where V(***K, X***) is presented for two time moments t1 and t2 in the input space ***X*** (X1, X2).

Thus its desirable state can be symbolically represented by "a ball" in Fig. 9b; and according to the forth doctrine, a new model of real systems' decentralized adaptive control principles and algorithms C_M and C_R must be created.

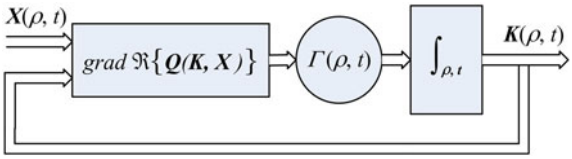


Fig. 10 The stochastic approximation performed according to the space-time gradient procedure

8 Stochastic Approximation Approach

According to [15, 31] and other researches, the branch of complex systems science—the GFSCT—implies the possibility to organize a successful decentralized adaptive control process aiming to improve "our reality". The backbone of such an approach is seen in the space-time extension of the well-known stochastic approximation procedure combined with the computerized fuzzy verbal and perceptual reasoning [32]. The idea to extend and use the stochastic approximation procedure is evoked by the fact that the potential function $\mathbf{V}(K, X)$ itself in our case cannot be determined explicitly in terms of classical mathematics. Instead, we are able to measure or evaluate some decisively important under investigation crisply or fuzzily. Usually it is required to minimize or maximize the averages of those characteristics $\mathcal{M}$; (here $\mathcal{M}$ means a mathematical expectancy). When substituting unknown potential V-function for the available set of characteristics $\mathbf{Q}$, it is convenient to use a certain additive or multiplicative function $\Re\{^*\}$ permitting to form only one $\Re\{\mathcal{M}\{\boldsymbol{Q}(\boldsymbol{K}, \boldsymbol{X})\}\} \rightarrow min$.

The convergence of $\mathbf{K}(\rho, t)$ towards $\mathrm{K_C}$ during the process of stochastic approximation can be performed according to the gradient procedure

$$\frac{d\boldsymbol{K}(\rho, t)}{d\rho} = \Gamma_\rho(\rho, t) \operatorname*{grad}_\rho \Re\{\boldsymbol{Q}(\boldsymbol{K}, \boldsymbol{X})\}$$

$$\frac{d\boldsymbol{K}(\rho, t)}{dt} = \Gamma_t(\rho, t) \operatorname*{grad}_t \Re\{\boldsymbol{Q}(\boldsymbol{K}, \boldsymbol{X})\}$$

symbolically performed by the algorithm shown in Fig. 10 [33–35].

It is well known [15, 32] that such a procedure converges only probabilistically:

$$P\left\{\lim_{\rho,t\rightarrow\infty} [\boldsymbol{K}(\rho, t) - \boldsymbol{K}_C] = 0\right\} = 1$$

9 Conclusion

The assessment of correspondence between systems and their models must be built on the comparison of causal information and knowledge base.

Two control loops must be involved in the process of reality and its model's improvement.

Four working doctrines serve as a backbone of newly born General Fuzzy Systems Control Theory (GFSCT).

Each socio-technical fabric should be analyzed in the independent space-time coordinate system having in mind that space is sometimes anisotropic.

Stochastic approximation procedures may serve as a tool for the minimization of a certain potential V-function as a universal principle of existence of each system in the real world.

References

1. Yew-Soon, O., Menghiot, L., Xianshun, C.: Memetic computation past, present & future. IEEE Comput. Intell. Mag. **5**(2), 24–31 (2010)
2. Peng, Q., Jane, W.Z., Ray Liu, K.J.: Genomic processing for cancer classification and prediction. IEEE Signal Process. Mag. **24**(1), 100–110 (2007)
3. Jasinevicius, R.: Why today's systems theory can't cope with global environmental or marine systems catastrophes and crises? In: Proceedings Baltic International Symposium (BALTIC), 2010 IEEE/OES US/EU, 24–27 Aug 2010-Riga, Latvia, p. 8 (2010)
4. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5621636
5. Hahn, H.-J. et al. (eds.): The Future of Democracy: An Indian Perspective, pp. 97–127. Verlag des Professorenforums, Giessen (2004)
6. Tackling Poverty: The Roles of Business, Government, and NGOs. In: Ethix magazine (Seattle, USA), Issue 50. http://www.ethix.org/article.php3?id=346. Accessed Nov 2006
7. Kyoto protocol to the United Nations framework convention on climate change. http://unfccc.int/resource/docs/convkp/kpeng.html
8. Kristopher, G., Andreas, L., Shane, S., Paul, W.: Making sense of the subprime crisis. In: Brookings Papers on Economic Activity, Fall 2008, Conference Draft. http://www.brookings.edu/economics/bpea/bpea.aspx
9. Anthony, B., Michael, O.: An introduction to evolutionary computation in finance. IEEE Comput. Intell. Mag. **3**(4), 42–55 (2008)
10. Krugman, P.: How did Economists Get it so Wrong? The New York Times, 6 Sept 2009, p. 14 (2009)
11. Jasinevičius, R., Petrauskas, V.: Nonlinear and dynamic extensions for fuzzy cognitive maps (FCM) tools. In: Information Technologies' 2009: 15th International Conference on Information and Software Technologies, IT 2009, Kaunas, Lithuania, 23–24 April 2009, pp. 11–15 (2009)
12. Jasinevičius, R., Krušinskienė, R., Petrauskas, V., Tkaciov, A.: Dynamic fuzzy expert maps: idea and implementation. In: Information Technologies' 2011: Proceedings of the 17th International Conference on Information and Software Technologies, IT 2011, Kaunas, Lithuania, 27–29 April 2011, pp. 17–22 (2011)
13. Wiener, N.: Cybernetics or Control and Communication in the Animal and the Machine, 2nd edn, p. 212. The MIT Press, Cambridge/Wiley and Sons, New York (1961)
14. Hussein, A.A., Sameer, A., Axel, B.: MEBRA: multiobjective evolutionary-based risk assessment. IEEE Comput. Intell. Mag. **4**(3), 29–36 (2009)
15. Collective Adaptive Systems. In: Expert Consultation Workshop 3 & 4 Nov 2009. Report. European Commission, Information Society and Media, p. 17 (Nov 2009)
16. Aizerman, M.A., Braverman, E.M., Rozonoer, L.I.: Theoretical foundations of the potential function method in pattern recognition learning. Autom. Remote Control **25**, 821–837 (1964)
17. Jasinevicius, R.: European roadmap for complex systems science. In: Information Technologies' 2011: Proceedings of the 17th International Conference on Information and Software Technologies, IT 2011, Kaunas, Lithuania, 27–29 April 2011, p. 15 (2011)

18. Pollock, N.: Knowledge Management and Information Technology, p. 384. Defense Acquisition University Press, Fort Belvoir, Virginia (2002)
19. Ahsan, S., Shah, A.: Data, information, knowledge, wisdom: a doubly linked chain? http://ww1.ucmss.com/books/LFS/CSREA2006/IKE4628.pdf
20. Čenytė, J., Jasinevičius, R.: Apie kontekstinio panašumo mata. Informacinės technologijos : 16-oji tarpuniversitetinė magistrantu ir doktorantu konferencija : konferencijos pranešimu medžiaga / Kauno technologijos universitetas, Vytauto Didžiojo universitetas, Vilniaus universiteto Kauno humanitarinis fakultetas. Kaunas : Technologija. ISSN 2029–249X. 2011 (in Lithuanian), p. 133–136 (2011)
21. Zadeh, L.A.: Toward human level machine intelligence-is it achievable? The need for a paradigm shift. IEEE Comput. Intell. Mag. **3**(3), 11–22 (2008)
22. Kacprzyk, J., Zadrozny, S.: Computing with words is an implementable paradigm: fuzzy queries. Linguistic data summaries, and natural- language generation. IEEE Trans. Fuzzy Syst. **18**(3), 461–472 (2010)
23. Richard, C.D., Robert, H.B., Modern Control Systems, p. 1018. Prentice Hall, Saddle River (2008)
24. Klir, G.: Architecture of Systems Problem Solving, p. 354. Plenum Press, New York (1985)
25. Gupta, V. H., Wagener, T., Liu, Y.: Reconciling theory with observations: elements of a diagnostic approach to model evaluation. Hydrol. Process. **22**, 3802–3813 (2008)
26. Mendel, M.J., Wu, D. Perceptual Computing: Aiding People in Making Subjective Judgments, p. 320. Wiley, Hoboken (2010)
27. Wunch, G.: Systemtheorie, Leipzig. Akademische Verlagsgesellschaft Geest & Portig K.-G., p. 240 (1974)
28. Chen, M., Trefethen, A., Banare-Alcantara, B., Jirotka, M., Coecke, B., Ertl, T., Schmidt, A.: From data analysis and visualization to causality discovery. IEEE Comput. **44**(10), 84–87 (2011)
29. Jasinevičius, R., Petrauskas, V.: Dynamic SWOT analysis as a tool for environmentalists. In: Environmental Research, Engineering and Management, Kaunas: Technologija, No 1(43), p. 14–20 (2008)
30. Gardner, M.R., Ashby, W.R.: Connectance of large, dynamical (cybernetic) systems: critical values for stability. Nature **228**, 784 (1970)
31. Jasinevichius, R.: Parallel space-time structure for computer vision systems. Informatica **3**(3), 418–431 (1992)
32. Kang, J.M., Kwon, J.K.: The East Asian Model of Economic Development, No.25(Jung_Mo_Kang).pdf, pp. 1–21. http://web.ias.tokushima-u.ac.jp/naito/No.25(Jung_Mo_Kang).pdf (2011)
33. Borkar, V.S.: Stochastic approximation: a dynamical systems viewpoint. Tata Institute of Fundamental Research, Mumbai, p. 172. http://www.tcs.tifr.res.in/~borkar/trimROOT.pdf (2008)
34. Jasinevicius, R.: Parallel Space-Time Computing Structures. Mokslas, Vilnius (in Russian), p. 183 (1988)
35. Dongrui, W., Mendel, M.J.: Perceptual reasoning for perceptual computing: a similarity-based approach. IEEE Trans. Fuzzy Syst. **17**(6), 1397–1411 (2009)
36. Tsypkin, Y.Z.: Adaptation and Learning in Automatic Systems, p. 291. Academic Press, Newyork (1971)

Emergent Phenomena in Natural Complex Systems

Jiri Bila

Abstract In this chapter are described some properties of emergent phenomena and situations that appear in natural complex systems. There are introduced three classes of unexpected situations and two of them belong to emergent situations. For the detecting possibility of appearance of emergent situation is used the indication of violation of structural invariant of the complex system. Here is used only one type of structural invariant—matroid and matroid bases. A simple calculus for the emergent situation appearance computation is introduced. The application of the presented approach and computation method is demonstrated by three simple emergent situations: the change of the strategy in swarm colony, traffic jam and floods.

Keywords Emergent situations · Structural invariants · Matroid bases · Ramsey numbers

1 Introduction

After an enormous effort of physicists to discover beginnings of nature phenomena (as some typical nature shapes) and especially of the origin of our Universe (as the essential and largest emergence) it was formulated a vision of mathematic formalisms for future with the following properties:

- To be more qualitative than quantitative.
- To contain some traces of emergent phenomena that lead to nature shapes [1].
- To allow the direct work with shapes as with qualitative entities, not with their descriptions (e.g., by analytical mathematical descriptions).

J. Bila (✉)
Department of Instrumentation and Control Engineering, Faculty of Mechanical Engineering, CTU in Prague, 160 00 Prague 6, Czech Republic
e-mail: bila@vc.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_9,

Author [2] titled such a discipline *Morphomatics*.

Emergent phenomena for the sake of their uniqueness form obstacles to a traditional approach in a search for solutions of a given task. It is possible to name long file of interesting books and chapters that stroked a non perforable wall.

In this chapter we slightly shifted this wall though not too much. Application of the approach that has been formed for emergent situations in role of natural shapes as a consequences of violation of so called *Structural invariants*, e.g., in [3] and continued, e.g., in [4], has been oriented in last years to ecosystems [5, 6] and [7]. In this chapter are introduced formal and computational tools from recent works in more general context.

2 Some Related Works

Great interest in detecting emergent situations, understood in a general sense, has appeared in the field of fault diagnosis [8–11]. However, emergent situations are also of interest in other fields. There are important emergent situations in safety engineering and in control of complex systems [12–15].
Essential knowledge sources about complex systems, emergence phenomena and complexity are concentrated in [16–18]. Very original and ambitious seems to be, e.g., the interpretation of complexity as a linguistic variable [17].

A great effort in specification of the essence of emergent phenomena was recognized in [19, 20] but also in [4, 3] and [21]. Especially—first two works, though very different in environments, noted relations between the nature of emergent phenomena and the sign system in which the emergent phenomena are represented.

Comparing approaches presented in works introduced above with our approach presented in this chapter there is possible to find a few differences (explained in further sections). Here we preliminarily introduce only two of them:

- An emergent phenomenon in a system is indicated (detected) in our case by the violence of Structural Invariants and is not necessarily related to some critical quantities or limits of complexity of the system.
- The appearance of an emergent situation is represented in our case more as a qualitative and discrete structure than as reaching of some value interval in a preformed scale.

3 Types of Emergent Situations

Many definitions of emergence phenomena and of emergent situations were introduced in past time and we do not like to contribute with another definition. In this section will be introduced properties of so called surprising situations and two of them will be considered as emergent situations.

3.1 Situations that were Surprising and in a Certain Context can be Considered as Surprising (A)

Their causes and their output forms (outputs, shapes) are known. It is possible to recognize them and predict their appearance. Examples of processes and systems that generate such situations are, e.g.: Belousov-Zhabotinski reaction; environments for initiation of solitons; oregonator; brusselator. *They all belong to the field of Synergetics.* Such situations are **not** Emergent Situations.

3.2 Situations that were Surprising and are Still Surprising (B)

Their causes are not known however their output forms are known. Such situations are **Emergent situations** (EMSs) . They have the following properties:

(b1) The situation appears suddenly without explicit association with situations of the previous relevant context in the system.
(b2) The situation appears as a discrete object, fact, shape.
(b3) The global reason of such a situation appearance is a violation of the system structure (not of the system function).
(b4) The detailed reasons and the internal causes of the appearance of the situation are not known (i.e., it is impossible to propose a complete model of situation evolution and prediction).
(However the shape of situation has been recognized – i.e., is known how such a situation looks.)
(b5) The appearance of such a situation is possible to detect.

Situations that belong to this class are, e.g.: change of strategy of behavior in a swarm colony; appearance of floods; appearance of rough waves; traffic jams.

3.3 Situations that will be Surprising (C)

Neither their causes nor their output forms are known. Such situations are **Emergent situations** (EMSs). They have the following properties:

(c1) The situation appears suddenly without explicit association with situations of the previous relevant context in the system.
(c2) The situation is assumed and surmised as a discrete object, fact, shape.
(c3) The global reason of such a situation appearance is a violation of the system structure (not of the system function).
(c4) No model of such a situation is available before it first operate.
(c5) The appearance of such a situation is possible to detect.

Situations that belong to this class are, e.g.: Possible instabilities in Ecosystems; sharp growth of extents of large towns; appearance of artifacts in nano-structures; situations of discoveries in Conceptual Design; the violation of supersymmetries in quantum mechanics.

4 Cognitive Tools for Investigation of Conditions of Appearance of Emergent Situations (CAEMSs)

In a classical approach to modeling we search for some measurable representative variables and we try to determine and compute some limit values indicating *Conditions of Appearance of Emergent Situations* (CAEMSs).

For example in case of a sudden appearance of traffic jam we find variables as

⟨ the density of car file, the average velocity of cars, the average distance between cars, the average number of braking of individual car during half of an hour, the average number of stopping of individual car during half of an hour ⟩

and we verify by simulation the variants of values of these variables—here is one of them:

⟨ the density of car file (150 cars /km), the average velocity of cars (80 km/hour), the average distance between cars (10 m), the average number of braking of individual car during half of an hour (15), the average number of stopping of individual car during half of an hour (7) ⟩.

This approach is very "economic" however it does not represent relations between individuals. (It is inconvenient, e.g., for the monitoring of flights.) For rather multi-dimensional systems we have to descend more deep into principles of modeling:

4.1 Level of the Description (LD)

There is distinguished between external and internal LD of a system.
External description

- External description is usually introduced by a set of external variables that may characterize the system, emergent phenomena and emergent situations. Such external variables are associated with human experience with behavior of the considered system and emergent phenomena (in case of EMSs of the type B), e.g., the density and velocity of cars in a traffic jam) and with type and behavior of the environment in which we expect emergent phenomena (in case of EMSs of the type C), e.g., the evapotranspiration of plants or loss of biodiversity in an ecosystem).

Internal description

- Constructed internal description could allow to represent interactions between elements (components) of the system.

- In constructed LD can be used a few types of modeling schemes, e.g., "Model of Reality (M_R)", "Sign Model—Interpretation Space (SM/IS)", "Macrostructure—Microstructure (MA/MI)", as it was described, e.g., in [3].
- LD sharply distinguishes between model of structural features and model of functional properties.

4.2 Structural Invariants (SIs)

- SIs "defend" the structure of the system against possible transformations.
- SIs represent a "constant part" of interactions between elements (components) of the system.
- Violation of SIs indicates and detects possibility of an appearance of EMS.

In [3], there were introduced and described some essential types of SIs, e.g., "Matroid and its Bases", the pair "Dulmage-Mendelsohn Decomposition, Tree Ordering", the pair "Hasse Diagram, Set of Associated Rules" and "Algebra of Transformations on the Set of Situations". In this chapter will be used only one type structural invariant: *Matroid and matroid bases***.**

4.2.1 Matroid and Matroid bases as a Structural Invariants of a Complex System

Matroid has the following pleasant properties:

- It is possible to construct it for each set of elements when we have the relation of independence or when they are given independent sets.
- There are investigated all elements with regard to relation of independence.
- The relation of independence is very adaptable (from unary, binary till n-ry relation) with necessary semantic contents.

Matroid is usually introduced as the following structure

$$M = \langle X, IND, \{N_1, N_2, ..., N_n\}\rangle, \tag{1}$$

where X is the ground set of elements (components), *IND* is a relation of independence and $N_1, N_2, ..., N_n$ are independent sets. Matroid *bases* $(\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_m)$ are maximum (according to cardinality (#)) independence sets.
Matroid is considered and constructed (in our case) on a "controlled volume" of the complex system (in the next text called "basic group of elements"). This basic group is represented as a ground set (X) of the matroid.

In case that relation *IND* is considered as a *binary relation* it is possible to use following consequences:

- The independent sets on X are discovered as perfect sub-graphs (in a perfect graph on X).
- The cardinalities of node sets of these perfect sub-graphs have limits in [0, #X], (where [x, y] denotes (and will denote) the interval of integers between x and y including x,y).
- For the modeling of the discoveries of bases ($\boldsymbol{B}_1, \boldsymbol{B}_2, \ldots, \boldsymbol{B}_m$) the formalism of Ramsey numbers—R(#$\boldsymbol{B}_i$, #N_j), i ∈ [1, m] and j ∈ [1, n] is used.

Note 3.1: Till now there are precisely known only some Ramsey numbers (RNs), e.g.: R(3, 3) = 6, R(3, 4) = 9, R(3, 5) = 14, R(3, 6) = 18, .., and for others were computed only intervals (sometimes rather wide) R(3, 15) = [73, 78], …, R(4, 4) = 18, R(4, 11) = [96, 191], …, R(6, 10) = [177, 1171], …, R(10, 10) = [798, 23 556], …, R(19, 19) ≥ 17 885. (In computations in Sect. 5 will be used only known table quantities.)
Hypothesis 3.1: Possible *appearance* of EMSs in a modeled complex system is detected by violence of a structural invariant—in our case—of one of the matroid bases. As the violence of the matroid basis is considered (in our case) the extension of the basis by (at least) one element.
Note 3.2: The extension of a matroid basis by (at least) one element may be attained already by extension of X by a minimum number of elements needed for the achievement of the nearest higher Ramsey number.

Example 3.1: #X = 1600. (#$\boldsymbol{B}$ = 11 for #X ≥ 1597) and for one element extension (#$\boldsymbol{B}$ = 12 for #X ≥ 1637) it is necessary to add at least 40 elements.

Example 3.2: The extension of basis B by one element is illustrated in Fig. 1. The perfect graph in Fig. 1 has six nodes and 15 (brown) edges. Coloring the edges by green and blue colors, there appears at least one perfect sub-graph with 3 nodes and 3 edges (basis $\boldsymbol{B}$) — for example the green one—R(3, 3) = 6. For extension of $\boldsymbol{B}$ by one element (into $\boldsymbol{B} + \boldsymbol{1}$ with 4 nodes) we need to add at least 3 elements (that are invisible here) — R(3, 4) = 9.

Note 3.3: Steps of the work with matroid:

S1. The relation of independence is introduced.
S2. The matroid is constructed.
S3. The bases on the matroid are extracted.
S4. The extension of some basis is executed.

4.2.2 Reinterpretation of Some Results of Relativistic Theory of Information

One of the essential results of Relativistic theory of information [22] were the equations for the computation of the conditional entropy of a system. These equations were invariant respect to Lorentz transformation. We use here only the structure of those equations.
Complexity of the matroid basis extended by one element $H_{COM}(\boldsymbol{B+1})$

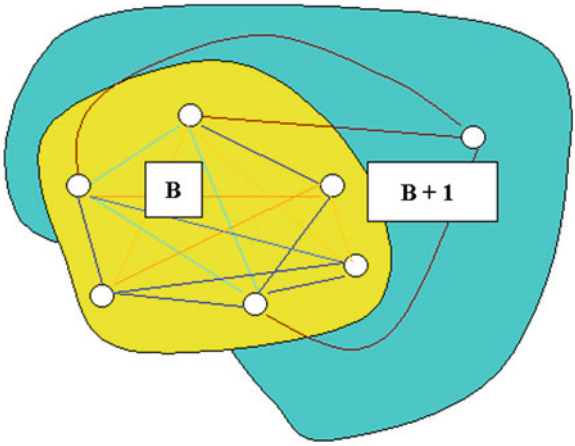


Fig. 1 Extension of a matroid basis by one element

$$H_{COM}(\boldsymbol{B+1}) = H_{COM}(\boldsymbol{B}) + u H_P(\boldsymbol{B}), \tag{2}$$

Power of the emergent phenomenon modeled by one element extension of the basis $\boldsymbol{B}$.

$$H_P(\boldsymbol{B+1}) = H_P(\boldsymbol{B}) + (u/c) H_{COM}(\boldsymbol{B}), \tag{3}$$

Contribution of the power of the emergent phenomenon

$$\Delta H_P(\boldsymbol{B+1}) = (u/c) H_{COM}(\boldsymbol{B}), \tag{4}$$

where u is the quotient of self-organization $u \in \langle 0, c \rangle$, c is the limit of self-organization (both depends on the emergent environment). $H_{COM}(.)$ is approximated in our case by number of elements of matroid basis, $\Delta H_P(\boldsymbol{B+1})$ is the needed power of the emergent phenomenon expressed in percentage (for example, contribution for 20 % is calculated as (120/100) = 1.2) .

4.2.3 Computation with $H_{COM}(B)$, $H_P(B)$ and $\Delta H_P(B+1)$

C1. The quantity of $\Delta H_P(\boldsymbol{B+1})$ for a given emergent environment is estimated.
C2. The quantity of *(u/c)* for a given emergent environment is determined.
C3. The number of elements of Basis $\boldsymbol{B}$ is computed.

$$H_{COM}(\boldsymbol{B}) = \xi((c/u)\Delta H_P(\boldsymbol{B+1})) = \#\boldsymbol{B}, \tag{5}$$

where ξ (x) is the nearest higher complete number (e.g.,ξ (2.5) = 3).

C4. The number (#*X*) of interacted elements needed for emergent situation appearance is computed.

Example 3.3: For $\Delta H_P(\boldsymbol{B+1}) = 2$ and (u/c) = 0.5 is computed $\#\boldsymbol{B} = 4$ and Ramsey numbers: 9, 18, 25, 41, 61, …, 282. For the extension of the basis by one element (#(***B+1***) = 5) we find Ramsey numbers: 14, 25, 49, 87, …, 464. Pairs of Ramsey numbers with minimum added elements needed for *induction of emergent situation* are (9→14/5), (18→25/7), (25→25/0), (41→49/8), (61→87/26), etc.

4.2.4 Estimation $\Delta H_P(B+1)$ and Computation of (u/c)

Estimation of $\Delta H_P(\boldsymbol{B+1})$ is done by means of quantities of external variables x_i, i = 1, …, n estimated for emergent (x_{iem}) and for nominal (x_{inom}) situations:

$$\Delta H_P(\boldsymbol{B+1}) = (\Sigma(\omega_i(x_{iem} - x_{inom})/x_{inom})^2)^{1/2}, \text{ for } i = 1, \ldots, n, \tag{6}$$

where ω_i are quotients of importance (computed by Saaty method [23]).
The quantity of (u/c) is computed by iteration procedure using relations derived from the equation of information homogeneity:

$$I_{i1}{}^*N_{i1} = I_{i2}{}^*N_{i2} = I_{i3}{}^*N_{i3} = \ldots = I_{im}{}^*N_{im} = I_S{}^*1, \tag{7}$$

where I_{ij} for j = 1, …, m are quantities of information needed for the solution of a problem by N_{ij} steps. I_S is a super quantity of information by which is the problem solved in one step.
The quantity of (u/c) is computed by iteration procedure (avoiding the necessity to know quantities I_{ij}).

$$(u/c) = (I_{ij}/I_S) = (1/k)(1/N_{ij}), \text{ for some "i, j" adequate to solved problem.} \tag{8}$$

Where k is a normalization constant, $k \in \langle 0, 1\rangle$. As a solved problem is considered the induction of connectivity between the interacted elements from the set X (the ground set of elements of Matroid) and N_{ij} is the number of bonds between elements on X. Though we do not know neither #X nor the scheme of connectivity on X, the *maximum* of bonds between elements on X is

$$N_{ij} = (\#X)(\#X - 1)/2. \tag{9}$$

The iteration procedure starts with the quantity $\Delta H_P(B + 1)$ computed by external variables and with some given quantity of $(u/c)_0$. Then is computed the number of elements in the matroid basis $\#\boldsymbol{B}$, from tables is selected a lowest Ramsey number R($\#\boldsymbol{B}$, Y), (Y is an integer from the table) and there is computed the number N_{ij} by (9):

$$N_{ij} = (R(\#\boldsymbol{B}, Y))(R(\#\boldsymbol{B}, Y)-1)/2. \quad (10)$$

Then is computed $(u/c)_1$ with given normalization constant k

$$(u/c)_1 = (I_{ij}/I_S) = (1/k)(1/N_{ij}). \quad (11)$$

In case when the absolute normalized value of the difference is higher than 0.15

$$\delta = |((u/c)_1 - (u/c)_0)/(u/c)_0| \geq 0.15, \quad (12)$$

the procedure is repeated with higher $(u/c)_0$(for the case $u/c)_1 < (u/c)_0$) and v.v.

5 Simple Examples of Appearance of Emergent Situations

5.1 Change of the Strategy of Behavior in a Swarm Colony

The changes of termite behavior (e.g., the formation of attack patterns) are induced by the changes of termite pheromone *concentration*.

There exists a basic group of termites (the Basis) that causes the higher increase of pheromone concentration and induces the spreading of message for other termites. Using the calculus that had been explained in Sect. 4 we may compute the number of termites from basic group and the number of termites needed for the emergence of other type of behavior.

For $\Delta H_P(\boldsymbol{B+1}) = 1.5\, and\ for\ (u/c) = 0.2$ (lower self-organization)

$$H_{COM}(\boldsymbol{B}) = \xi((c/u)\Delta H_P(\boldsymbol{B+1})) = \#\boldsymbol{B} = 8.$$

For $\#\,\boldsymbol{B} = 8$ there are RNs: 28, 56, 84, 127, 495, 1031, 1870.

For $\#(\boldsymbol{B+1}) = 9$ (the extension by one element) there are RNs: 36, 69, 121, 316, 1713, 565, 153, 780, 6588.

Pairs of RNs for *induction of an emergent situation* are (28→36/8), (56→69/17), (127→153/26), (84→115/31), (495→565/70).

5.2 Traffic Jams

Traffic jam is an emergent result of interaction of many transport elements and factors (cars, lights, structure of transport symbols, …, whether, hours in the day time, etc.) and of an "self-organizing phenomenon" that forms the external performance of

traffic jam. Some problems from this field were published in [24]. Here is considered only the emergent phenomenon that starts in a basic group of cars.

"Self-organization" process is hard to describe but is possible to compute numbers of elements in basic groups:

$$\text{For } \Delta H_P(\boldsymbol{B+1}) = 2 \, and \, for \, (u/c) = 0.4$$

$$H_{COM}(\boldsymbol{B}) = \xi((c/u)\Delta H_P(\boldsymbol{B+1})) = \#\boldsymbol{B} = 5.$$

For $\# \boldsymbol{B} = 5$ there are Ramsey Numbers: 9, 18, 25, 35, 41, 49, 61, 153, 198, 230, 242, 282, 417.

For $\#(\boldsymbol{B+1}) = 6$ (the extension by one element) there are RNs: 18, 35, 41, 58, 87, 102, 165, 298, 374, 434, 495, 548, 614, 710, 780, 878, 1070, 1171.

Pairs of RNs for induction of an emergent situation are (9→18/9), (25→35/10), (35→41/6), (49→58/9), (282→298/16), (417→434/17).

5.3 Floods

The floods are consequences not only of an enormous rainfall, accumulation and motion of water but also of the changes of behavior of brooks, streams, rivers, ponds and lakes. The induction of emergent situation "floods" ("conspiracy" of water carriers) starts on a basic group of components. Though a needed self-organization process is difficult to discover and describe, the numbers of components in basic groups is possible to estimate.

$$\text{For of } \Delta H_P(\boldsymbol{B+1}) = 1.5 \, and \, for \, (u/c) = 0.09,$$

$$H_{COM}(\boldsymbol{B}) = \xi((c/u)\Delta H_P(\boldsymbol{B+1})) = \#\boldsymbol{B} = 17.$$

For $\# \boldsymbol{B} = 17$ there are Ramsey Numbers: 92, 182, 284, 548, 627, 737, 8917.

For $\#(\boldsymbol{B+1}) = 18$ (the extension by one element) there are RNs: 98, 614, 722, 871, 11005.

Pairs of RNs for induction of an emergent situation are (92→98/6), (548→614/66), (627→722/95), (737→871/134).

6 Conclusions

A general conclusion that has been achieved in this chapter may by considered, from the first sight, trivial: "By increasing the concentration of mutually interacting elements (or their properties) in a finite space may be induced an emergent phenomena."

However—not in every case are induced emergent phenomena as a consequence of increasing density of interacting elements in closed space. It depends on self organization phenomenon that had to be realized and also on the secret of volume of the basic group. This second factor we tried to catch in this chapter.

Second problem—the possibility of an emergent situation appearance is detected however we do not know if it will look as we assume. E.g., as a result of training of Kohonen neural network we discover some clusters that seem to be meaningful [20]. Well—the clustering has emerged but in the semantic content of the clusters we could be absolutely confused. In other words—the explanation and semantic description of the discovered EMS belong to another sophisticated process—Interpretation. (It holds —unfortunately —also for our three examples from Sect. 5).

In this chapter has been proposed a procedure that (after tuning) could compute something as "a distance" from appearance of a possible emergent phenomenon. This computation has been performed in this chapter only for emergent situation for which are known consequences and not causes (type B).

For emergent situations of the type C we have no models because these situations are available not soon then after their first operation. Such situations are possible only to detect by special methods. A technique of matroid bases and Ramsey numbers implicitly explains an apparent stability of ecosystems though just in ecosystems are expected emergent situations of the type C (and some ecological catastrophes demonstrated it).

Acknowledgments The development of this chapter has been supported by Research Grant SGS12/177/OHK2/3T/12. This support is very gratefully acknowledged.

References

1. Reid, R.G.B.: Biological Emergences. Evolution by Natural Experiment. A Bradford Book. The MIT Press, Cambridge (2007)
2. Stewart, I.: Nature's Numbers (The Unreal Reality of Mathematical Imagination). In: Brockman, J. (ed.) The Orion Publishing Group, Great Britain (1995)
3. Bila, J.: The detection of emergent situations by structural invariants. In: 14th International Conference on Soft Computing—Mendel 2011, pp. 534–539. Brno, June 2011
4. Bila, J., Gojda, S.: Monitoring emergent situations in complex systems with the Help of conceptual detection patterns. Int. Rev. Autom. Control **4**, 855–866 (2011)
5. Bila, J., Pokorny, J., Jura, J., Bukovsky, I.: Qualitative modelling and monitoring of selected ecosystem functions. Ecol. Model. **222**, 3640–3650 (2011)
6. Bila, J., Bukovsky, I.: Qualitative models for the landscape development monitoring. J. Commun. Comput. **9**(6), 721–728 (2012)
7. Bila, J. and Pokorny, J.: Design of smart regions and landscapes, Invited paper. In: 30th cCAADe Conference, Prague (2012)
8. Abramovici, M., Breuer, M.A. and Friedman, A.D.: Digital System Testing and Testable Design, pp. 99–104. IEEE Press, New York (1995)
9. Chess, B.: Accounting for the unexpected: fault diagnosis out of the ivory tower. In: International Test Conference, pp. 1135–1142. D (1998)

10. Wang, H., Wu, Q.H.: Detect and diagnose unexpected changes in the output probability density function for dynamic stochastic system: an identification approach In: 14th World Congress of IFAC, vol. 17, pp. 217–222. Beijing (1999)
11. Yan, Z., Lam, J., Huijun, G.: Fault detection for fuzzy systems with intermittent measurements. IEEE Trans. Fuzzy Syst. **17**, 398–410 (2008)
12. Atkins, E.M., Durfee, E.H. and Shin, K.G.: Expecting the unexpected: detecting and reacting to unplanned-for world states. In: 13th National Conference on Artificial Intelligence, p. 1377. Portland (1996)
13. Besnard, D. and Greathead, D.: A cognitive approach to safe violations. Cogn. Tech. Work **5**, 272–282 (2003)
14. Mourão, H. and Antunes, P.: A collaborative framework for unexpected exception handling. In: 11th International Workshop CRIWG 2005, pp. 168–183. Brazil, September 2005
15. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
16. Zelinka, I., Sanayei, A., Zenil, H., Rössler, O.E.: Emergence Complexity and Computation in Nature. Springer (2013)
17. Sanayei, A.: Complexity as a linguistic variable. Complex Syst. **20**(3), 253–264 (2012)
18. Zelinka, I., Skanderova, L., Davendra, D.D. et al.: Controlling complexity, In: Numerical Analysis and Applied Mathematics—ICNAAM 2012, pp. 654–657. Kos, (2012)
19. Darley, V.: Emergent phenomena and compexity. In: Fourth International Workshop on the Synthesis and Simulation of Living Systems—Artificial Life IV, pp. 411–416. MIT Press, Cambridge (1996)
20. Ultsch, A.: Emergence in self organizing feature maps, In: 6th International Workshop on Self-Organizing Maps—WSOM 2007, 2007. http://bielcoll.ub.uni-bielefeld.de
21. Bila, J.: Algebras of transformations in the detection of unexpected situations of UX^3 type. In: 13th International Conference on Soft Computing—Mendel 2010, pp. 495–501. Brno (2010)
22. Jumarie, G.: New results in relativistic information theory. Application to deterministic, stochastic and biological systems. Int. J. Syst. Sci. **7**(4), 393–414 (1976)
23. Saaty, T.L.: Exploring the interface between hierarchies. Multiple objectives and fuzzy sets. Fuzzy Sets Syst. **1**, 57–68 (1978)
24. Nagel, K., Rasmussen, S.: Traffic at the edge of chaos. In: 4th International Workshop on the Synthesis and Simulation of Living Systems, MIT (1994)

Evolutionary Systems in Complex Signal Analysis

Tomas Brandejsky

Abstract All complex systems presenting chaotic behaviour are non-linear ones and many problems of their analysis and modelling are caused by application of linear or pseudo-linear models which are not able to represent all aspects of signals generated by these systems. Experiments with some natural-based signal data like e.g. EEG ones concluded presence of typical composite periodic functions, like sin(sin(x)). These functions have specific behaviours which will be presented. Especially, they are non-stationar, have continuous spectrum and thus it is hard to apply usual tools like Fourier transform. To analyse these signals, evolutionary GPA-ES system was used.

Keywords Complex system signals · Signal analysis · Evolutionary systems · Generic programming algorithm

1 Introduction

Many complex systems producing chaotic behaviour are extremely complicated in comparison to simplest deterministic chaotic systems used frequently as examples of chaotic systems. These systems contains from hundreds (e.g. transportation systems) to hundreds of billions of nonlinear elements (cells in the brain). There it is hard or even impossible to model and analyse properties and behaviours of such systems analytically and it is need to constrain our models to input—output models on the place of state space representation. The initial analysis of EEG signals was published in the work [1]. This work pointed out the analysis of EEG signal by GPA-ES evolutionary system [2]. Used EEG data were measured during car driver micro sleep experiments [3, 4] and they were used in raw form without any pre-processing like

T. Brandejsky (✉)
FTSci CTU in Prague, Konviktska 20, 110 00 Prague 1, Czech Republic
e-mail: brandejsky@fd.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_10,

Table 1 List of the most frequent algebraic structures discovered in EEG signal models

```
Y(t) = (c1 + t) Sin[c2 + t] Sin[(c3 + t + Sin[t])Sin[c4 + t]]
Y(t) = c1 t(c2 + c3 Sin[c4 t]) + c5 t Sin[c6 + c7 t]
Y(t) = (c1 + t) (c2 + t) + c3 (c4 + t) Sin[Sin[(c5 + c6t)]]
Y(t) = (c1 + t + c2 t) (c3 + c4 t + c5 Sin[t] + Sin[c6 t])
Y(t) = c1 + (c2 + t) (c3 + c4 t) + c5 Sin[(c6 + c7t)]
Y(t) = (c1 + t + c2 Sin[c3 t]) Sin[c4 Sin[c5 + Sin[c6 t]]]
Y(t) = Sin[c1 t] (c2 + Sin[c3 + c4 t])
Y(t) = c1 + c2 t Sin[c3 t] Sin[c4 (c5 + t)]
Y(t) = c1 + Sin[c2 t] + Sin[c3 Sin[Sin[c4 Sin[t (c5 + t) + c6]]]]
Y(t) = c1 + Sin[Sin[c2]Sin[c3] Sin[c4 t]] + Sin[c5 Sin[c6 t]]
Y(t) = c1 + t + c2 (Sin[c3 t] + Sin[t + c4 t]) + Sin[c5 + c6 (c7 + c8 t)]
Y(t) = c1 + t^2 + (c2c3 t) (c4(c5 + t) + Sin[c6 + t]) Sin[c7 (c8 + t)]
       Sin[c9 (c10 + t)]
Y(t) = c1(c2 + Sin[c3(c4 + t)])Sin[Sin[Sin[c5(c6 + t)
       (t + Sin[c7 + t])]]]
Y(t) = c1 + c2 Sin[c3 t] + Sin[c4 + t]
Y(t) = (c1 + c2 t) (t (c3 + t) + Sin[c4 t]) + Sin[c5 Sin[c6 t]]
```

especially filtering. Without any apriory information, symbolic regression by evolutionary system produced local models of the signal containing algebraic structures listed in the Table 1. On the base of these results, it is possible to form nonlinear GPA building blocks (named user defined functions by Koza in [5]) based on the structures frequently identified in the signal (1–3), where this list is constrained to different forms of composite functions of two sin() functions only. The graphs of these functions are displayed as Fig. 1. Discovered functions are periodic one, but their properties are complex in comparison to their short algebraic form.

$$y = \sin\left(c1^* \sin\left(c2^* t + c3\right) + c4\right) \tag{1}$$

$$y = \sin\left(c1^* t + c2\right)^* \sin\left(c3^* t + c4\right) \tag{2}$$

$$y = \sin\left(c1^* t + c2^* \sin\left(c3^* t + c4\right) + c5\right) \tag{3}$$

2 Composite Functions Behaviours

Composite functions in common, not only the above presented ones have many specific features in contrary to simple functions. Inherited complexity of the above discovered functions (1–3) might be easily illustrated by their Taylor series, derivations (and impossibility to integrate them symbolically—the common feature of composite functions).

While Taylor series of sin(x) (4) and sin(sin(x)) (5) are similar except higher significance of the next elements, the symbolic representations of Taylor series of

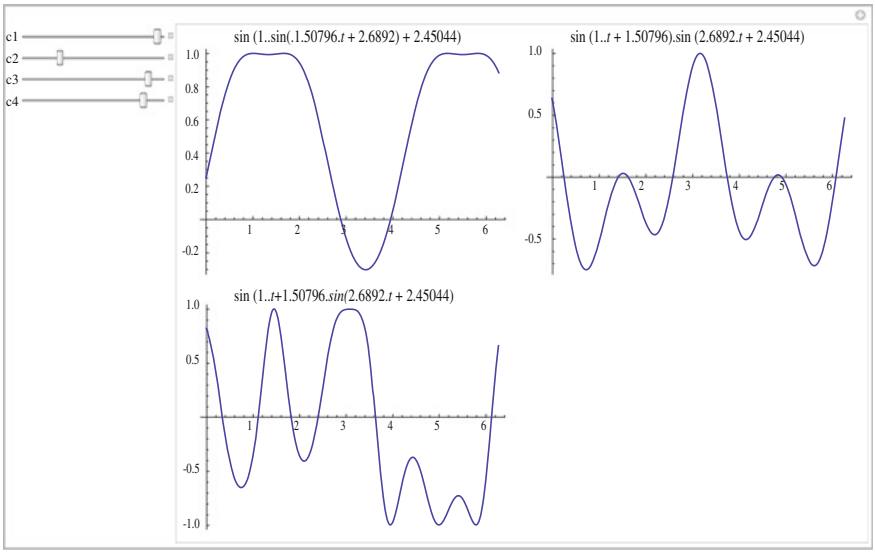


Fig. 1 Examples of functions (1–3) for printed parameters

$\sin(c_1^* x + c_2)$ (6) and $\sin(c_1^* \sin(c_2^* x + c_3) + c_4)$ (7) strongly differs and the form of the expression (7) points to extremely complicated nature of this expression.

$$y = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \frac{x^9}{362880} + o(x)^{11} \tag{4}$$

$$y = x - \frac{x^3}{3} + \frac{x^5}{10} - \frac{8x^7}{315} + \frac{13x^9}{2520} + o(x)^{11} \tag{5}$$

$$y = \sin(c_2) + c_1 \cos(c_2)\, x - \frac{1}{2}\left(c_1^2 \sin(c_2)\right) x^2 - \frac{1}{6}\left(c_1^3 \cos(c_2)\right) x^3 + o(x)^4 \tag{6}$$

$$\begin{aligned} y = {} & \sin(c_1 \sin(c_3) + c_4) + c_1 c_2 x \cos(c_3) \cos(c_1 \sin(c_3) + c_4) \\ & + x_2\left(-\frac{1}{2} c_1^2 c_2^2 \cos^2(c_3) \sin(c_1 \sin(c_3) + c_4)\right. \\ & \left. - \frac{1}{2} c_1 c_2^2 \sin(c_3) \cos(c_1 \sin(c_3) + c_4)\right) \\ & + \frac{1}{6} x^3 \left(-c_1^3 c_2^3 \cos^3(c_3) \cos(c_1 \sin(c_3) + c_4)\right. \\ & + 3 c_1^2 c_2^3 \sin(c_3) \cos(c_3) \sin(c_1 \sin(c_3) + c_4) \\ & \left. - c_1 c_2^3 \cos(c_3) \cos(c_1 \sin(c_3) + c_4)\right) \\ & + o(x)^4 \end{aligned} \tag{7}$$

Taylor series of expressions (2) and (3) tends to even more complicated results than (7). The complexity of composed functions projects itself into derivations too. The complexity of many studied signals like EEG one increases with derivation degree as well as complexity of derivations of above mentioned composed functions.

The first and second derivations of (1) are (8) or (9) respectively. Analogically derivations of (2) are (10 and 11), derivations of (3) are (12 and 13).

$$y' = c_1 c_2 \cos(c_2 t + c_3) \cos(c_1 \sin(c_2 t + c_3) + c_4) \tag{8}$$

$$y'' = -c_1^2 c_2^2 \cos^2(c_2 t + c_3) \sin(c_1 \sin(c_2 t + c_3) + c_4) - c_1 c_2^2 \sin(c_2 t + c_3) \cos(c)_1 \sin(c_2 t + c_3) + c_4) \tag{9}$$

$$y' = c_3 \sin(c_1 t + c_2) \cos(c_3 t + c_4) + c_1 \cos(c_1 t + c_2) \sin(c_3 t + c_4) \tag{10}$$

$$y'' = c_1^2(-\sin(c_1 t + c_2)) \sin(c_3 t + c_4) - c_3^2 \sin(c_1 t + c_2) \sin(c_3 t + c_4) + 2 c_1 c_3 \cos(c_1 t + c_2) \cos(c_3 t + c_4) \tag{11}$$

$$y' = (c_1 + c_2 c_3 \cos(c_3 t + c_4)) \cos(c_1 t + c_2 \sin(c_3 t + c_4)) \tag{12}$$

$$y'' = -c_2 c_3^2 \sin(c_3 t + c_4) \cos(c_1 t + c_2 \sin(c_3 t + c_4)) - \sin(c_1 t + c_2 \sin(c_3 t + c_4))(c_1 + c_2 c_3 \cos(c_3 t + c_4))^2 \tag{13}$$

As expected, the complexity of derivations of composite functions (1–3) increases in contradictory to polynomial or periodic ones. But integration brings problems too. There does not exist algebraic way to compute indefinite integral of these functions, there is possible to compute numerical solution of definite one. These numerical solutions of integrals (1–3) are displayed on Figs. 2, 3 and 4 respectively.

The above presented functions discovered by evolutionary algorithm in the EEG data are capable do describe this signal. Because they are not oriented to signal component frequencies but also to speed of its change, they reflect different view on signal than Fourier transform. It is also need to comment the physical sense of identified components (1–3).

(1) represents frequency modulation,
(2) describes amplifier fading up sin(t) signal with variable reinforcement following the other sin(t) component
(3) is similar to (1) with added constant (within the given time window) increasing or decreasing of the signal frequency.

3 Evolutionary Algorithm in the Complex Signal Analysis

Complex signal analysis like e.g. above discussed analysis of EEG one brings specific problems given by analysed data amount. The studied data requires processing of data samples of EEG measured in 30 channels. The data were collected for 24 h continuously with sampling frequency is 64 Hz. Developed models for each channel

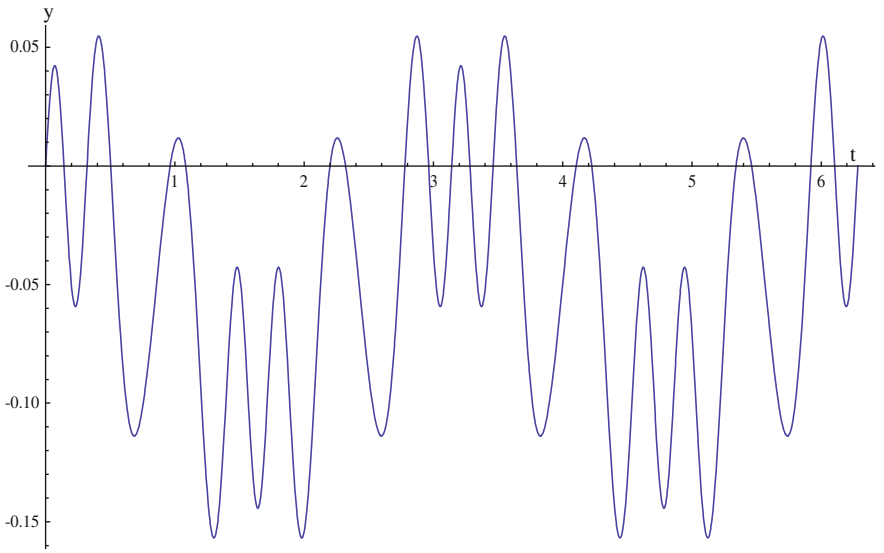


Fig. 2 Example of definite integral of $y = \sin(c_1 \sin(c_2 t + c_3) + c_4)$ function for $c1 = 10$, $c2 = 2$, $c3 = 3$ and $c4 = 0$

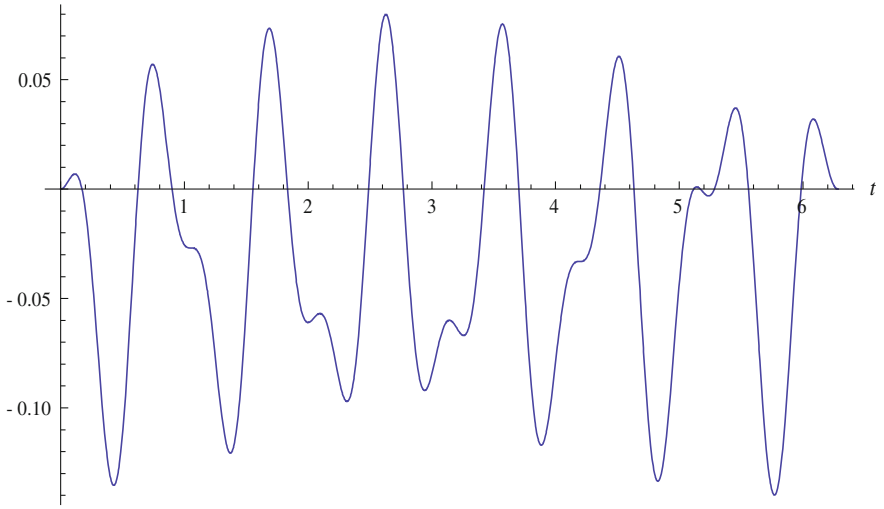


Fig. 3 Example of definite integral of $y = \sin(c_1 t + c_2)\sin(c_3 t) + c_4)$ function for $c1 = 10$, $c2 = 2$, $c3 = 3$ and $c4 = 0$

requires identification of many coefficient too, because regressed models cannot be restricted to second order ones. Analogical situation exists in the analysis of transportation data, where is the lower sampling frequency, but the number of channels sometimes creates equal data flow.

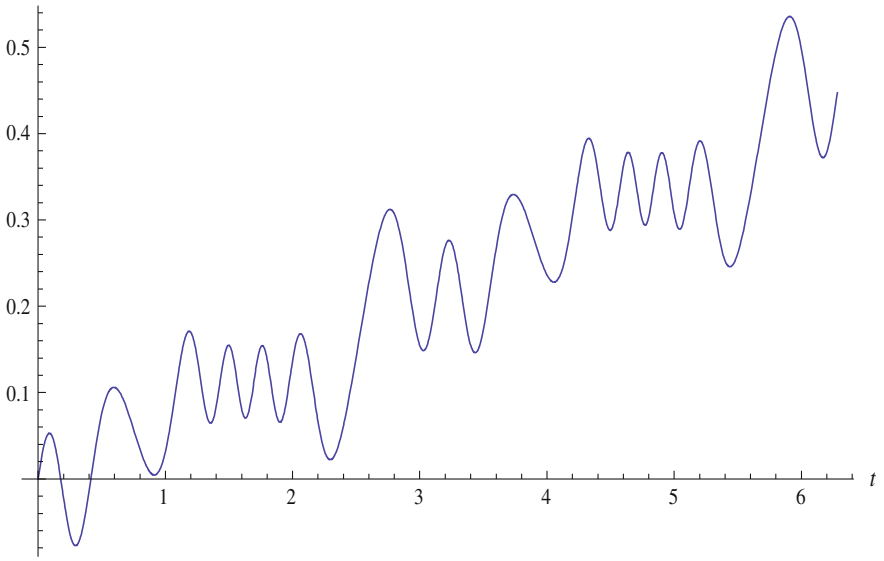


Fig. 4 Example of definite integral of $y = \sin(c_1 t + c_2 \sin(c_3 t + c_4) + c_5)$ function for $c1 = 4$, $c2 = 10$, $c3 = 2$, $c4 = 3$ and $c5 = 0$

The situation tends to numerically intensive computations and requires parallel implementation of the used algorithms. The used GPA-ES evolutionary system works efficiently on multiprocessor computers with shared memory (UMA architecture) and it exists in implementation using OpenMP library. The implementation in the Chapel language was tested too [6], but the OpenMP implementation is more efficient now.

In constrained available HW resource environment, there is need to maximize efficiency of the used algorithms. Thus, in the past time the main effort was focussed to optimize the GPA-ES. There were studied questions of optimal population sizes including extremely small ones [7] , the problem of optimal building block set was studied in the work [8] from the viewpoint of chaotic system symbolic regression.

Data are divided into small series of 200–300 samples by sliding window technique. For each series, the local model is formed. This technique is applied for observed non stationarity of the signal and it allows producing of simple models because it is not need to form models of parameter change in the time.

The structure of the used algorithm is displayed on Fig. 5. As it was discussed in the paper [1], there are many ways of fitness function evaluation. For the solved problems of signal analysis, each variable is regressed independently on the others as a function of time. Errors were measured in each time step where the prediction was calculated from the starting point given by measured data—thus prediction horizon was restricted to single time-step. This approach gives higher structural sensitivity than the others discussed in [1].

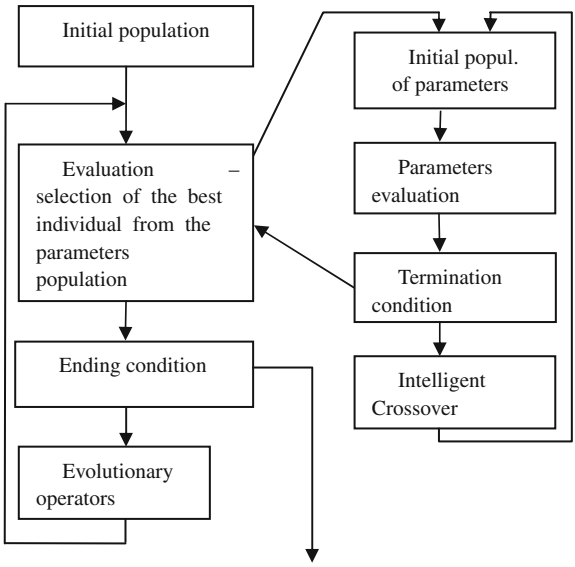


Fig. 5 The structure of hybrid GPA-ES algorithm with GPA (*left* column) individual parameters optimization by inherited ES (*right* column of blocks)

4 Conclusions

Evolutionary system GPA-ES allowed to discover presence of composite functions in EEG signal and in the transportation data measured in the Prague city. This algorithm was used without apriory information about expected form of the data symbolic regression and the obtained results were presented in the Table 1. These results conclude presence of the signals which it is hard to analyse by standard techniques like FFT and the need of development of novel form of analysis respecting composite functions.

Acknowledgments This work was supported by the research project of MŠMT ČR No 6840770043 "Improvement of methods of design and employment of transportation networks from optimization viewpoint".

References

1. Brandejsky, T.: The use of local models optimized by genetic programming algorithm in biomedical-signal analysis. In: Handbook of optimization From Classical to Modern Approach, pp. 697–716. Springer, Heidelberg (2012). ISBN 978-3-642-30503-0
2. Brandejsky, T.: Multi-layered evolutionary system suitable to symbolic model regression. In: Recent Researches in Applied Informatics, vol. 1, pp. 222–225. WSEAS Press, Athens (2011)
3. Faber, J., Pekny, J., Pieknik, R., et al.: Simultaneous recording of electric and metabolic brain activity. Neural Netw. World **20**(4), 539–557 (2010)

4. Bouchner, P., Faber, J., Novotny, S., Tichy, T.: Driver‘s attention level improvement with use of biofeedback stimulation incorporated into driving simulator. Neural Netw. World **19**(1), 109–118 (2009). ISSN 1210–0552
5. Koza, J.R., Bennett III, F.H., Andre, D., Keane, M.A.: Genetic Programming III: Darwinian Invention and Problem Solving. Morgan Kaufmann, San Francisco (1999)
6. Brandejsky, T.: Parallel implementations of GPA-ES algorithm. In: Mendel 2012. Brno: VUT in Brno, Faculty of Mechanical Engineering, pp. 30–34 (2012). ISBN 978-80-214-4540-6
7. Brandejsky, T.: Small populations in GPA-ES algorithm. In: Mendel 2013. Brno: VUT in Brno, Faculty of Mechanical Engineering, pp. 31–36 (2013). ISBN 978-80-214-4755-4
8. Brandejsky, T.: Influence of operator set to chaotic system symbolic regression. In: Mendel 2013. Brno: VUT in Brno, Faculty of Mechanical Engineering, pp. 63–68 (2013). ISBN 978-80-214-4755-4

Macroscopic Description of Complex Self-Organizing System: Belousov–Zhabotinsky Reaction

Anna Zhyrova, Dalibor Stys and Petr Cisar

Abstract Based on the information theory of mutifractal objects was developed the method for analysis of complex self-organized system, such as living cells. To demonstrate some of the features of the analysis we choose the simplest system—the Belousov–Zhabotinsky reaction (chemical clock). It is always composed of observed sequence of states stable for certain period of time and the experimenter has full control of mechanical constraints imposed on the system. We use the Renyi information entropy equation for calculation of information gain by which a point contributes to the total information in the image. In this way we create characteristic vector of the system state in phenomenological coordinates of phase space. We have also derived related variables, the point information gain entropy and point information gain entropy density. The later values are unique to structured information. The ultimate goal of the method is to determine the characteristics of a system which best characterize momentary multifractal properties of the system. The relation between the phenomenological phase space and the internal coordinates of the system remain unknown.

Keywords Information entropy · State trajectory · Multivariate analysis

1 Introduction

Complex macroscopic structure of self-organized systems arises from a coherent interaction among lower level components of the system. Systems are sensitive to initial conditions. In experimental systems there is often observed hierarchical scaling which is often difficult to include into the measurement conditions where the

A. Zhyrova (✉) · D. Stys · P. Cisar
Institute of Complex Systems, Faculty of Fishery and Water Protection,
University of South Bohemia, Zámek 136, 37333 Nové Hrady, Czech Republic
e-mail: zhyrova@frov.jcu.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*,
Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_11,

resolution and precision is given by the measurement method. In the work with image records besides optical limits [1] we have to take into account aspects associated with image digitization (digital camera sensor imperfection; random noise; disadvantages of the algorithms, are using for debayer transformation, demosaicing (parser matrix) and compression of the final image, etc. [2–4]). A long list of problems of representation of the three dimensional scene by the two dimensional matrix of image points may be prepared. To solve these problems, we propose the method [5], which assesses the information yield from a given image while representing the characteristic structure of the object.

Another problem for experiment data interpretation is the complex unpredictable behavior of the examined objects, which are present self-organizing systems, which could be consider from different point of view [6, 7]. As experimental tool for analysis of the behavior of the self-organizing system we have chosen the Belousov-Zhabotinsky reaction [8, 9]. The distinct feature of the BZ reaction is that although it in fact consists of 80 chemical reactions, relatively simple patterns arise and may be modeled by approaches such as cellular automata. Most reports on BZ reactions come from experiments done in homogenous systems, either in mixed cuvettes or various types of flat reactors. The analysis does not require elaborate reconstruction of series of 3D images as in the case of bird flocks or fish schools, organ behavior or living cells. Moreover, the experimenter in case of chemical clock has full control of mechanical constraints imposed on the system.

2 Material and Methods

2.1 Belousov–Zhabotinsky Reaction

Experiments were performed with the oscillating bromated-ferroin-bromomalonic acid reaction (kit were provided by Dr. Jack Cohen) [10]. The reaction mixture was composed out of following solutions: 0.34 M sodium bromate, 0.2 M sulphuric acid, 0.057 M sodium bromide, 0.11 M malonic acid as substrate and redox indicator 0.12 M 1,10 phenantroline ferrous complex. All reagents were coherently mixed under temperature 22 °C and added into Petri dish.

2.2 Image Recording Conditions

The images were captured by Nikon D90 camera [11] in regime Time lapse shooting in interval 10 s between snapshots with Exposure compensation +1/3, ISO 200, Aperture f/18 and Shutter speed 1/10 s. For preventing the reflection on the surface of the reaction mixture for photographing was used diffused light. To avoid the inevitable blurring the smallest detail of the image caused by compression using the

JPEG algorithm all the resulting images saved in NEF format provided by camera manufactured company with further proceeding into 12 bit TIFF format by ViewNX (Nikon software).

2.3 Calculation of the Image Information Yield

As it were shown in the work [12], multifractal structure of the object can be expressed in terms of Renyi information entropy through the parameter α. Inner structure of the multifractal object, which is related to the distribution of the physical, chemical or other quantities of system and thus provided their geometric properties, could be expressed through their singularities [13]. According this statement, the system phase transition could be described in term of Renyi entropy. However, we do not actually observe value of internal orthogonal variable but their projection into observable, phenomenological, variables. Thus, later in the text we assume that the space of phenomenological variables is multifractal equally well as its generating chaotic attractor.

We calculate the Renyi entropy for the data set containing the examined point and the dataset in which the examined point was excluded. The difference between these two values Point Information Gain (γ_α(x, y)) for given entropy of the order α:

$$\gamma_\alpha(x, y) = \frac{1}{1-\alpha} ln\left(\sum_{i=1}^{n} p_{i,x,y}^{\alpha}\right) - \frac{1}{1-\alpha} ln\left(\sum_{i=1}^{n} p_i^{\alpha}\right) \tag{1}$$

where $\rho_{i,x,y}$ and ρ_i are probabilities of occurrence of given intensity for given point x, y coordinate of camera pixel at given α in the image without and with the examined point. We do not used any algorithm to optimize α values, we just use a set of values in the range from 0,1 to 4.

Γ_α(i) is i-th observed value of γ_α (x, y) without specification of the point position and it is a unique characteristic for determining the system state into the terms of entropy value (Fig. 1). The question remains how to define the dataset used for calculation of the γ_α (x, y) should be, in structured objects the context of the location may not be overlooked.

In the next step number of points of given intensity is summed, we obtain Point Information Gain Entropy (H_α):

$$\mathbf{H}_\alpha = \sum_{x=1}^{x=n} \sum_{y=1}^{y=m} \gamma_\alpha(x, y) \tag{2}$$

For given dataset, i.e. image, the ordered tuple of dimensionless α coefficient and H_α is a unique characteristic of the image, i.e. each two different images will have different sets—to the extent of the precision of the digital imaging.

And on the final step of image processing were calculated Point Information Gain Entropy Density (Ξ_α):

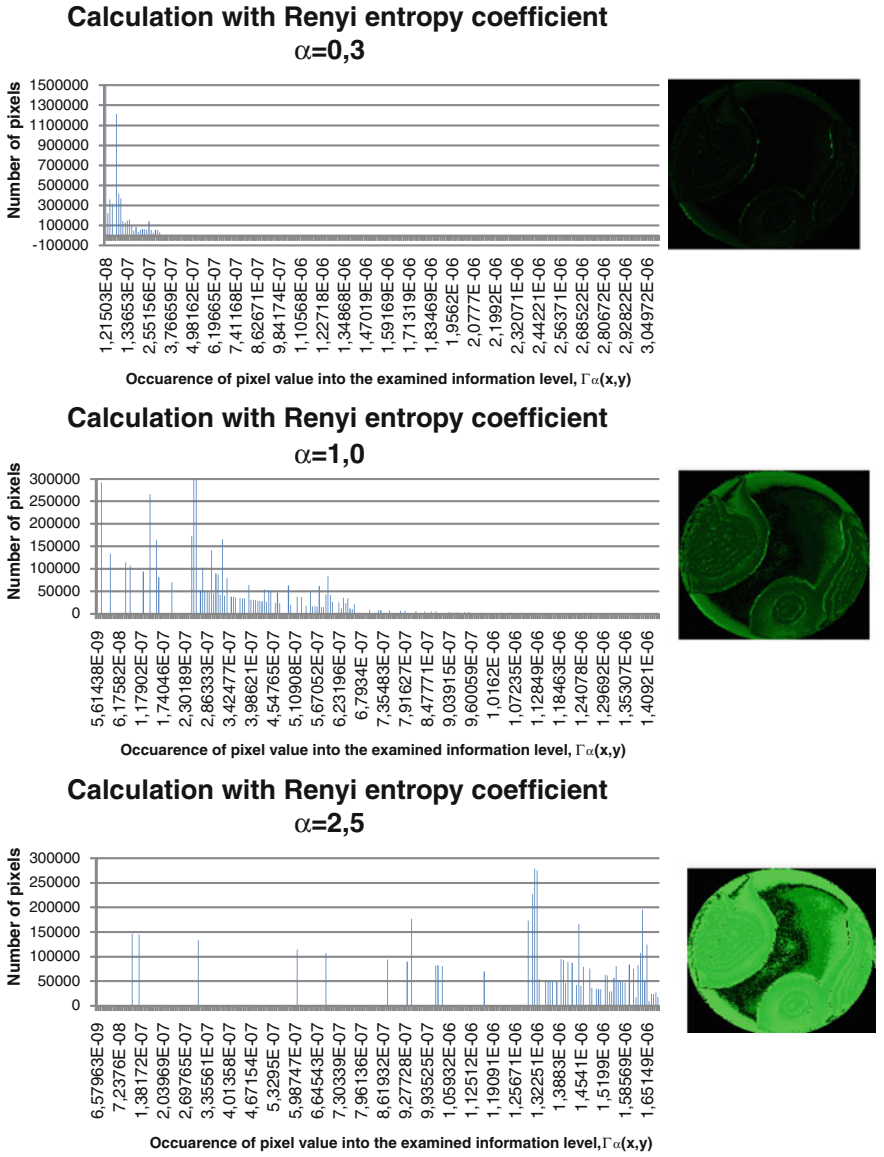


Fig. 1 The image represents Γ_α spectrum by which each particular point contributes to overall information in the image of the reaction dynamics. The *green channel* carries less noise signal than others [16] and we use it as example of the $\gamma_\alpha(x, y)$ transformed image. In the *left column* are depicted histograms of Γ_α -occurrence at given bit level. Rising the α coefficient changes the distribution of Γ_α levels. Large areas, with low Γ_α, represent the noise, medium Γ_α highlights edges of the image and finally rare, high entropy, points, represent *stripes* of chemically different regions in the medium

$$\Xi_\alpha = \sum_{i=1}^{i=n} \Gamma_{\alpha,i}\,(x,\, y) \tag{3}$$

Based on the central limit theorem, we expect that the resulting dataset is a subject to a normal distribution, we can apply the methods of multivariate analysis for drawing state trajectory in the face space for our developing system. The assumption of normal distribution is clearly not justified. Also, H_α values, which we use as coordinates, are linearly independent. Quite the contrary, they more probably are not orthogonal for many reasons both fundamental and technical. Using statistical approaches, such as principal component analysis (PCA) we may construct orthogonal spaces which best fit the observed dataset. Instead of constructing the state function as (unknown) manifold in a given coordinate system, we search coordinate system in which the state function is a plane of multidimensional space. Also the statistical distribution is not known and thus we replace it by the normal distribution. If we stick just with the experimentally observed data, we may base the interpretation on the work of Zampa [14]. In his general system theory he proposes to separate trajectory into segments which are within the measurement limits in part Markovian. Contemporary theories of the chaotic dynamic [15] show how the trajectory segments arise from chaotic dynamics. This gives us a reason to separate the individual phases of the reaction development on the next stage of multivariate analysis by means of cluster analysis.

2.4 Software Packages

For calculation of $\gamma_\alpha(x, y)$, H_α and Ξ_α values were used software packages developed by the Institute of Complex Systems. Stable versions of this software are available at http://www.expertomica.eu/software.php and are constantly updated. Latest versions are available upon request. The principal component analysis was performed using Unscrambler software provided by CAMO Company [17].

3 Results and Discussion

The use of H_α for describing the evolution of the system as a multifractal object using the values as its resulted in logical sequence of clusters in the new orthogonal, although still phenomenological, state space. The definition of point information gain as local variable also enables to demonstrate how a specific information contribution is distributed and how it contributes to the value of Rényi entropy, respectively H_α and Ξ_α. At the Fig. 1 it is shown that the α lower than one highlights the less probable objects (unique points, borders of object), where α higher than one highlight the large areas (background) and it should be chosen to emphasize just one particular distribution contributing to overall distribution in the image. On the Fig. 2 is shown decomposition of the system trajectory of the Belousow–Zhabotinsky reaction into

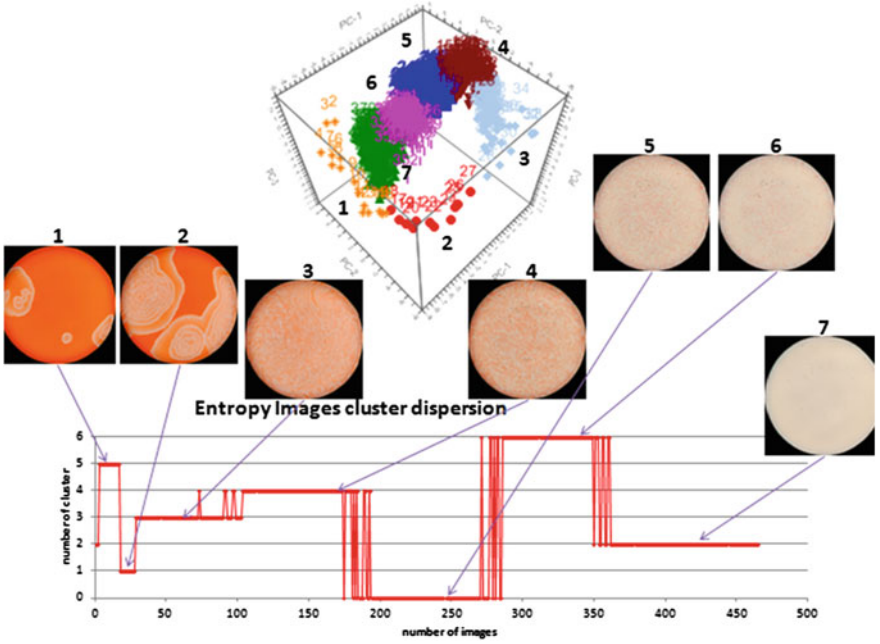


Fig. 2 The state trajectory of the BZ reaction performed in the Petri dish. Principle component analysis of the H_α allows to designate different states of the system, from the point of view of the method clusters of points in the state space. Scores differ significantly between clusters oriented and form logical trajectory in the principal coordinates space

series of states which are for distinct period of time asymptotically stable under current conditions. Clusters (series of images) are very well separated and consistent in time, for each of group could find characteristic image represented the state of the system in their developing process. Moreover, each of the clusters states should have its own spectrum of H_α and Ξ_α values which characterizes it.

Thus applying of information entropy as the basic characteristics of the image is a promising area for further research, with the ultimate aim of which is to create the reliable method of automated segmentation of the self-organizing system state space obtained by non-invasive imaging method. And the trajectory segmentation may be used despite to the fact that we do not know the proper manifold in the coordinate space.

Acknowledgments This work was supported and co-financed by the CENAKVA CZ.1.05/2.1.00/01.0024 and by the South Bohemia University grant GAJU 134/2013/Z.

References

1. Murphy, D.B.: Fundamentals of Light Microscopy and Electronic Imaging. Wiley-Liss, New York (2001)
2. Lorre, J.J., Gillespire, A.R.: Artifacts in digital imaging. Appli. Digital Img. Proc. Astron. **264**, 123–135 (1980)
3. Holst, G.C.: CCD Arrays, Cameras, and Displays, 2nd edn. JCD Publishing and SPIE Press, USA (1998, in press).
4. Janesick, J.R.: Scientific Charge-Coupled Devices. SPIE-The International Society for Optical Engineering, Bellingham (2001)
5. Stys, D., Jizba, P., Papacek, S., Nahlik, T., Cizar, P.: (2012) On measurement of internal variables of complex self-organized systems and their relation to multifractal spectra, Sixth International Workshop on Self-Organizing Systems, IWSOS 2012. Delft, The Netherlands (2012)
6. von Foerster, H.: On Self-Organizing Systems and Their Environments. Self-Organizing Systems. Pergamon Press, London (1960)
7. Camazine, S., Deneubourg, J.-L., Nigel, R.: Self-organization in Biological Systems. Princeton university press, Princeton (2002)
8. Belousov, B.: Periodic processes of malonic acid oxidation in a liquid phase. Biophysics **9**, 306–311 (1959)
9. Rovinsky, A., Zhabotinsky, A.: Mechanism and mathematical model of the oscillating bromated-ferroin-bromomalonic acid reaction. Phys. Chem. **88**, 6081–6084 (1984)
10. Cohen, J.: Belousov-Zhabotinski Reaction: Do-it-Yourself Kit. United Kingdom, c2009-01. http://drjackcohen.com/BZ01.htm
11. Nikon D90 image quality: Imaging Resource, c1998. http://www.imaging-resource.com/PRODS/D90/D90IMAGING.HTM. Accessed 1998
12. Jizba, P., Toshihico, A.: The world according to Renyi: thermodynamics of fractal systems. AIP Conf. Proc. **597**, 341–348 (2001)
13. Halsey, Thomas C., Jensen, Mogens H., Kadanoff, Leo P.: Fractal measures and their singularities: the characterization of strange sets. Phys. Rev. A. **33**, 1141–1151 (1986)
14. Zampa P., Arnost, R.: 4th WSEAS International Conference on Systems Theory and Society (WSEAS), 4th WSEAS Conference. Wisconsin, USA (2004).
15. Cvitanovi'c, P., Artuso, R., Mainieri, R., Tanner, G., Vattay, G.: Web book Chaos: Classical and Quantum. ChaosBook.org, Niels Bohr Institute, Copenhagen, Available on (2012)
16. Digital camera image noise: Sean McHugh, (ed.) Cambridge in colour, c2005. http://www.cambridgeincolour.com/. Accessed 2005
17. CAMO Company Software. Norway. http://www.camo.com/rt/Products/Unscrambler/unscrambler.html (1984)

Part II
Systemic Modeling

Classical Invariants in the Quantum Mechanics of Chaotic Systems

F. Borondo

Abstract The relevance of classical invariants in the quantization and dynamics of quantum systems is discussed. Special attention is paid to the influence of periodic orbits ("*scars*") and the associated homoclinic and heteroclinic orbits. As an illustration we present some results concerning the vibrational dynamics of the LiNC/LiCN isomerizing system.

Keywords Nonlinear dynamics · Quantum chaos · Scar theory · Vibrational molecular dynamics

1 Introduction

The pioneering work of Poincaré on the three body problem at the turn of the twentieth century unveiled the possibility of chaotic motion in dynamical systems, also showing the importance of invariants such as periodic orbits (POs) and its homoclinic and heteroclinic connections in the hierarchical organization of the associated complex tangle [1]. The Kolmogorov-Arnold-Moser (KAM) [2] and Poincaré-Birkhoff-Lewis (PBL) [3] theorems completed Poincaré's ideas on the complicated, yet understandable, phase structure of chaotic systems. The former gives conditions under which chaos is restricted in extent, indicating that sufficiently irrational tori resist small, smooth perturbations in Hamiltonian systems. The latter established that in the destroyed rational tori an even number of fixed points remain under perturbation. These fixed points are alternately stable (elliptic) and unstable (hyperbolic). Around each elliptic fixed point there is a simultaneous application of the PBL and KAM theorems, which leads to a self-similar structure in phase space on all scales. In a previous paper [4],

F. Borondo (✉)
Departamento de Química and Instituto de Ciencias Matemáticas (ICMAT), Universidad Autónoma de Madrid, CANTOBLANCO, 28049 Madrid, Spain
e-mail: f.borondo@uam.es

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_12,

we showed how the stability island structure derived from the PBL theorem is also mimicked in quantum mechanics by the zeros of the Husimi function [5, 6] of states avoiding crossing in energy levels correlation diagram versus perturbation. This is only one of many results pointing out to the importance of some classical invariants to fully understand the structure of the quantum theory of classically chaotic systems.

The correlation of classical invariants with the corresponding quantum structures became very early a topic of much interest, since it was central to the development of the new mechanics [7]. In the case of an N-dimensional system with regular motion, the classical dynamics is confined to invariant tori [2], and the connection with the quantum treatment is very well understood in terms of the Einstein-Brillouin-Keller quantization conditions [8]

$$\oint_{\mathcal{C}_j} \mathbf{P} \cdot d\mathbf{q} = \left(n_j + \frac{\alpha_j}{4}\right) \qquad n_j = 0, 1, 2, \ldots, \qquad j = 1, \ldots, N, \tag{1}$$

where according to Einstein's prescription [9] the action integrals have to be evaluated on the N topologically independent circles, $\mathcal{C}_j$, defining the tori, n_j are integer quantum numbers, and α_j are the topological phases given by the Maslov indices [10, 11]. Equation (1) represents a typical semiclassical expression, in which a quantum condition is imposed into purely classical information. For each energy allowed by (1) there exists an associated WKB (Wentzel-Kramers-Brillouin) wave function [12]

$$\psi_{WKB} = \sum_j A_j \, e^{iS_j(\mathbf{q},\mathbf{I})/\hbar}, \tag{2}$$

where the sum extends to all possible branches, $S_j(\mathbf{q}, \mathbf{I})$, of the Hamilton-Jacobi equation.

In the chaotic regime, tori start to break and the condition (1) cannot, in general, be applied. However, in 1967 Gutzwiller [13] proposed a general quantization scheme, valid both in the chaotic and regular regimes, based in the POs and associated magnitudes of the system as the supporting classical invariants for the semiclassical quantization. The method is based on the use of the quantum mechanical Green function, $G(\mathbf{q}'', \mathbf{q}', E)$, giving the probability amplitude for a particle with energy E making the path $\mathbf{q}' \to \mathbf{q}''$. This (complex) function has poles at exactly the positions of the eigenvalues in the energy axis, and the residues give the corresponding eigenfunctions. Gutzwiller made a semiclassical approximation to this function, expressed as a sum over all classical trajectories satisfying the above mentioned conditions. This summation is very hard to carry out, especially when the dynamics of the system are chaotic, and then one has to resort to the trace, $g(E) = \int d\mathbf{q} \; G(q, q, E)$. This simpler function gives only information on the energy levels. Using stationary phase arguments a semiclassical expression for g was obtained as

$$g_{SC} = \frac{1}{i\hbar} \sum_{PO} \frac{T_0}{2\sinh(\chi/2)} \exp\left[i\left(\frac{S}{\hbar} - \ell\frac{\pi}{2}\right)\right], \tag{3}$$

which is the celebrated Gutzwiller's Trace Formula, in which T_0 is the period, χ the stability exponent, and S the action integral of each PO in the summation [13].

Another dramatic influence of POs on the quantum dynamics of classically chaotic systems was first discussed by Heller in his seminal paper of 1984 [14]. This author coined the term "scar" to name an enhanced quantum density localization in some of the system eigenfunctions. The origin behind this unexpected phenomenon, namely the building up of probability due to constructive interference, are the recurrences caused by the dynamics along an unstable PO [15], and this idea has been ever since very fruitful [16] in the field of quantum chaos. Bogomolny later demonstrated that this extra density can be obtained in the $\hbar \to 0$ limit by averaging in configuration space groups of eigenfunctions in an energy window [17], and the corresponding phase space version was investigated by Berry [18]. Other interesting aspects of scarring, such as the role of recurrences taking place through homoclinic and heteroclinic quantized circuits [19–23], the influence of bifurcations (in systems with mixed dynamics) [24], the scarring of individual resonance eigenstates in open systems [25, 26], or relativistic scars [27] have also been considered. Scars have also been experimentally observed [28–30], and their relevance in growing fields, such as nanotechnology [31, 32], optical microcavitities [33–38], optical fibers [39, 40], or graphene sheets [41], or their interest in constructing efficient basis sets for chaotic eigenstate calculations [42] has been described in the literature.

2 Model

The system that we have chosen to study is the LiCN molecule [43], whose vibrations can be adequately studied with the two degrees of freedom Hamiltonian

$$H = \frac{P_R^2}{2\mu_1} + \frac{1}{2}\left(\frac{1}{\mu_1 R^2} + \frac{1}{\mu_2 r_e^2}\right) P_\theta^2 + V(R, \theta), \tag{4}$$

in which the NC distance is frozen at its equilibrium value of $r_e = 2.186$ a.u., R is the distance from the NC center of mass to the Li atom, and θ is the angle between these two vectors. P_R and P_θ are the associated conjugate momenta, and μ_1 and μ_2 are the Li–CN and C–N reduced masses, respectively. The potential energy surface, $V(R, \theta)$, has been taken from the literature [44], and it is shown in Fig. 1 in the form of a contours plot. As can be seen, it presents two minima at $\theta = 0$ and π, corresponding to the two stable linear isomers LiCN and LiNC, respectively. The minimum energy path, $R_e(\theta)$, connecting these two isomers has also been plotted in the figure as a dashed line.

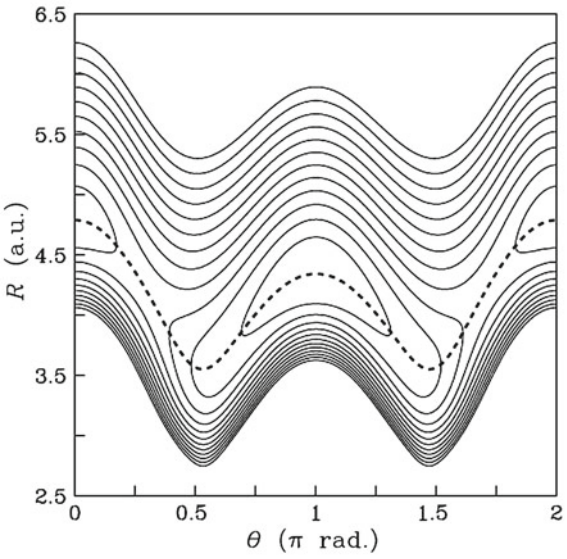


Fig. 1 Potential energy surface for the LiNC/LiCN isomerizing system. Contours lines have been plotted every 1000 cm^{-1}. The minimum energy path is shown as a *dashed line*

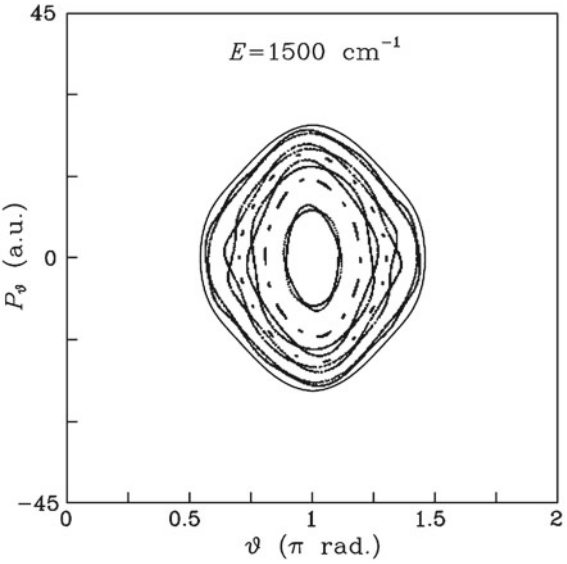


Fig. 2 Composite Poincaré surfaces of section for LiNC/LiCN at $E = 1500\,cm^{-1}$, showing the phase space structures (islands and bands of stochasticity) associated to resonances which are relevant to our work

3 Classical Calculations

The classical motion of the system can be monitored by means of Poincaré surfaces of section (SOS), which are best computed in the present case by taking the minimum energy path, $R_e(\theta)$, as the sectioning plane [45]. This SOS is made an area preserving map by defining a new set of canonical variables $\vartheta = \theta$ and $P_\vartheta = P_\theta - (dR_e/d\theta)P_R$.

The results corresponding to $E = 1500$ cm^{-1} are shown in Fig. 2. At this energy practically all phase space corresponds to regular motion confined into invariant tori. Chaotic dynamics are practically unnoticed, since restricted to extremely narrow bands along broken tori or the separatrices originated from resonant tori. Taking this into account, we have simplified the plot by including in it only the SOS associated to the lowest order resonant trajectories symmetrical with respect to the plane $\theta = \pi$ with ν_θ:$\nu_R = 1$:n. All possible POs of this type at this vibrational energy correspond to $n = 6, 7, 8, 9, 10$, and 10' and appear located consecutively in the plot from the center, $(\vartheta, P_\vartheta) = (\pi, 0)$ outwards [46–48]. Moreover, we have only represented in the n even cases the separatrices, which are enough to clearly sketch the shape and position of the chains of islands of the corresponding stable POs. For n odd we have chosen to represent the, in this case extremely narrow, chains of islands. Notice that here there appears a double number of islands than that indicated by index n, this being due to the fact that we have two possible senses in which the PO can be traced.

4 Quantum Calculations

For the quantum calculation of the vibrational states of LiNC/LiCN we have used the discrete variable representation method of Bacic and Light [49] with a final basis set of 2000 elements, enough to obtain the lowest 900 energy levels within an accuracy $\leq$0.01 cm^{-1}. The corresponding eigenstates, $|\mu\rangle$, in the range of energies considered in this paper are all regular, and can be classified according to the quantum numbers in the R and θ coordinates as (n_R, n_θ).

5 Quantum PBL Theorem

In this Section we will discuss the relationship existing between the PBL theorem and quantum mechanics. More precisely how the classical PBL theorem structures, visible for example in Fig. 2 for LiCN, can be seen in purely quantum data.

To unveil these structures in our system we use the correlation diagram for the vibrational quantum levels of LiNC/LiCN. As the variation parameter in the diagram we use $\hbar$. We artificially varied the value of this physical constant in order to change the number of quantum states that the regular region in the phase space energy shell can accommodate. Indeed, this number increases as smaller values of $\hbar$ are considered [50–52]. The results are shown in Fig. 3. Notice that we have chosen to represent in the vertical axis $E/\hbar$ instead of E. In this representation, the (mostly) harmonic states located in the (most stable) LiNC well render horizontal lines, while those corresponding to the LiCN isomer appear as hyperbolas, this shape arising due to the existence of the non-zero $\Delta E_{\text{LiCN-LiNC}}$ term corresponding to the energy difference between the two stable isomers of the molecule. Eye examination clearly reveals that below the line of circles, the levels simply interact in pairs at very sharp

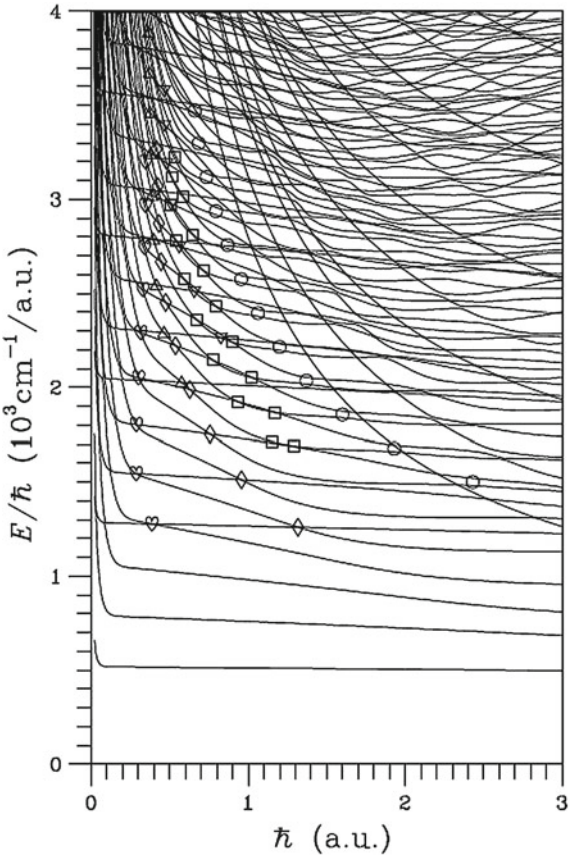


Fig. 3 Correlation diagram of vibrational energy levels for the LiNC/LiCN molecular system versus $\hbar$. Six different series of avoided crossings are indicated with different symbols corresponding to the quantum resonances: 1:6 (*hearts*), 2:14 (*triangles*), 1:8 (*diamonds*), 2:18 (*inverted triangles*), 1:10 and 1:10' (*squares*), and 1:8' (*circles*)

avoided crossings (ACs). Actually, the circles were found to constitute a threshold for the transition to (quantum) chaos, and also for the formation of scars, as discussed in Ref. [50, 51]. In this region there are six conspicuous series of sharp ACs, which have been marked with different symbols. Each of these series constitutes a family in which the wavefunctions that interact at the different ACs differ in the same number of quanta in both coordinates. For example, in the series marked with hearts, states avoiding crossing present a difference in the quantum numbers (see Sect. 4) of $|\Delta n_R| = 1$ and $|\Delta n_\theta| = 6$. In this way, states (0, 6) and (1, 0) interact in the lowest AC, the next AC the interaction is between states (0, 8) and (1, 2), and so on. It can be said that they are linked by a 1:6 quantum resonance.

Finally and more important, is that a careful examination of all ACs in the correlation diagram of Fig. 3 reveals the existence of a whole series of different families characterized by different values of $|\Delta n_R|$ and $|\Delta n_\theta|$. The result is summarized in Table 1. What it is extremely interesting is that all and each of the classical resonances existing in the phase space of the LiCN (which give rise to PBL island-separatrices

Table 1 Correspondence between classical resonances of LiCN in Fig. 2 and the series of avoided crossings in the correlation diagram of Fig. 3

Classical resonance in Fig. 2	Symbol in Fig. 3
1:6	♡
2:14	△
1:8	◇
2:18	▽
1:10	□
1:8'	◯

See text for details

structures; see Fig. 2) have a quantum counterpart as a family of ACs in the correlation diagram of vibrational levels of LiCN. This is, in our opinion, a very remarkable and important result in the field of quantum chaos.

Other aspects of this classical-quantum correlation effect has been considered elsewhere [4]. In particular, we have seen in the Harper map how the zeros of the Husimi quasiprobability density corresponding to the wavefunctions of the states involved in ACs mimic very precisely the shape and positions of the classical PBL resonance structures, and their evolution with the perturbation parameter as one sweeps along the AC, i.e. varies the value of the perturbation parameter.

6 Periodic Orbits, Homoclinic Circuits and Scars

As discussed in the Introduction, Poincaré discovered that POs were a key element in the hierarchical organization of complexity in the chaos showed by Hamiltonian systems. Also, much later Gutzwiller demonstrated that these orbits were also crucial for the quantization of such systems, especially when chaotic, and Heller showed how they influence or "scar" the quantum density of some eigenstates in this kind of systems.

In our group we have been working extensively in the computation of wave functions which are highly localized along unstable POs in classical chaotic systems. The essence of our procedure is well illustrated in the dynamical presentation of Ref. [53], where such functions were constructed by combining the ideas of Heller and Bogomolny on scars. The method consists in Fourier transforming for semiclassically quantized values of the energy, E_n, a wave packet, $|\phi(t)\rangle$, initially launched in the vicinity of the PO under study, after having being propagated for a short time $t \sim T = \mathcal{O}(T_{PO})$, in the following way

$$|\gamma_n\rangle = \frac{1}{2\pi\hbar}\int_{\infty}^{\infty} dt\; e^{-t^2/T^2} e^{iE_n t/\hbar}|\phi(t)\rangle. \tag{5}$$

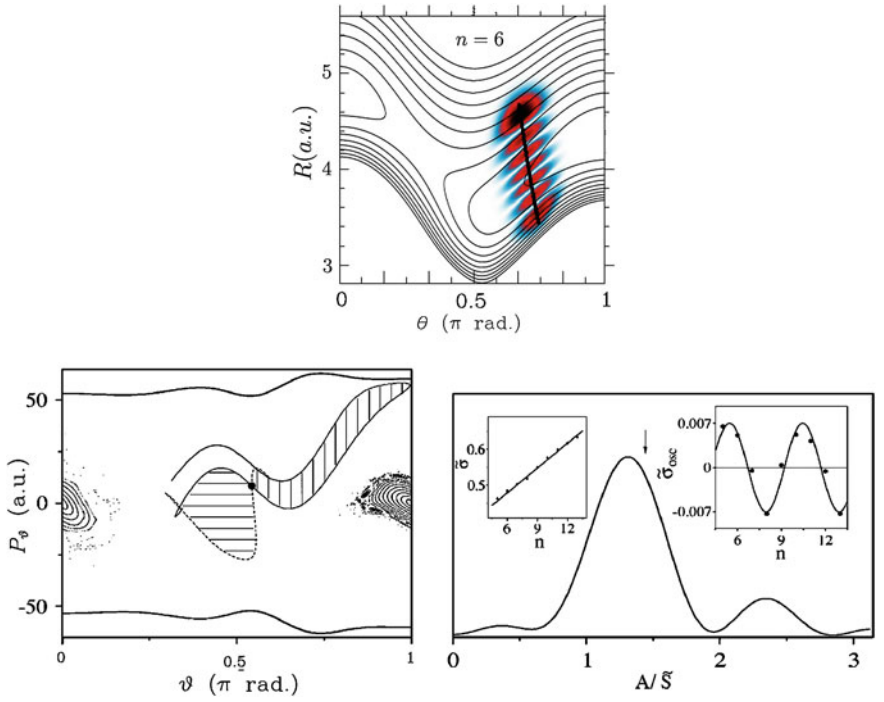


Fig. 4 (*Top*) *Squared* scar wavefunction computed with Eq. (5) for the saddle-node periodic orbit also represented in the plot. (*Bottom left*) Classical composite Poincaré surface of section for LiCN at $E = 9197.3\ \mathrm{cm}^{-1}$. Regular motion takes place at the invariant tori around LiCN and LiNC. The rest of the available phase space is left blank. Embedded in it, the fixed point (*full circle*) and unstable (*full line*) and stable (*dashed line*) manifolds corresponding to the PO in the top panel are represented. (Bottom right) Fourier spectrum of the oscillatory part (*right inset*) of the widths (6) made dimensionless (*left inset*) corresponding to the scar wave functions similar to those shown in the top panel for different values of n

(A more sophisticated construction of these functions can be found in Refs. [54, 55]). One example for a saddle-node bifurcation PO of LiCN [56–58], $|\gamma_6\rangle$, is shown in the upper part of Fig. 4. In the bottom left part of it, we also show the corresponding composite phase space picture at $E = 9197.3\ \mathrm{cm}^{-1}$, with the associated fixed point and manifolds and the regions of regularity around the two wells; the rest of the figure has been intentionally left blank to avoid visually complicating it. As can be seen the wave function, $|\gamma_6\rangle$, is highly localized along the PO and has six quanta of excitation.

These nonstationary functions, when projected on the eigenspectrum of the system, translate into localized bands centered at E_n, including the eigenstates within the corresponding widths [17]

$$\tilde{\sigma}_n = \sqrt{\sum_{\mu} \langle \mu | \gamma_n \rangle |^2 (E_\mu - E_n)^2}. \tag{6}$$

Surprisingly, these widths provides information on how the scarred structures are embedded in the quantum mechanics of the system. Actually, a Fourier analysis of their fluctuations around can be used to extract information about the relevant classical invariants. The corresponding results for $\tilde{\sigma}$ are shown with dots in the left inset on the right bottom panel of Fig. 4. When carefully examined, they are seen to oscillate around a straight line mean tendency. Elimination of this linear behavior renders the oscillatory part, $\tilde{\sigma}_{\rm osc}$, depicted in the right inset. Fourier transform of these values (main body of the panel) shows the existence of only a prominent peak in these oscillations centered at $S = 17.4$, value which is in very good agreement with the magnitude of the homoclinic phase $A_h = 20.1$ (arrow) computed by integration of any of the two shaded regions in the left panel of the figure. The interpretation of this result is clear, and has been thoroughly discussed in Ref. [21–23]. The analysis of the spectral characteristics of our scar wave functions requires, in addition, to the semiclassical quantization condition on the action of the PO (controlling the position of the corresponding peaks) another one consisting on the quantization of the phase space area of the "homoclinic torus" defined by the associated manifolds. When these two conditions are fulfilled simultaneously, scar states are best defined, and accordingly, its width in the eigenvalues spectrum is narrower. On the contrary, when only the first condition is fulfilled, scar states are more poorly described, and they appear more spread in the spectrum of eigenvalues.

Other classical invariants, such as the heteroclinic connections with a given PO or the Lazutkin invariant, have been shown by us to have an influence in the quantization of chaotic systems [21–23, 59].

References

1. Barrow-Green, J.: Poincaré and the Three Body Problem. American Mathematical Society, Providence (1997)
2. Arnold, V.I.: Mathematical Methods of Classical Mechanics. Springer, New York (1978)
3. Berry, M.V.: AIP Conf. Proc. **46**, 16 (1978)
4. Wisniacki, D.A., Saraceno, M., Arranz, F.J., Benito, R.M., Borondo, F.: Phys. Rev. E **84**, 026206 (2011)
5. Arranz, F.J., Seidel, L., Giralda, C.G., Benito, R.M., Borondo, F.: Phys. Rev. E **87**, 062901 (2013)
6. Leboeuf, P., Voros, A.: J. Phys. A **23**, 1765 (1990)
7. Jammer, M.: The Conceptual Development of Quantum Mechanics. American Institute of Physics, New York (1989)
8. Brack, M., Bhaduri, R.K.: Semiclassical Physics. Addison-Wesley, Reading (1997)
9. Einstein, A.: Ver. Phys. Ges. **19**, 82 (1917)
10. Maslov, V.: Théorie des Perturbations et des Méthodes Asymptotiques. Dunod, Paris (1972)
11. Maslov, V., Fedoriuk, M.V.: Semiclassical Approximations in Quantum Mechanics. Reidel, Boston (1972)

12. Landau, L.D., Lifshitz, M.E.: Quantum Mechanics: Nonrelativistic Theory. Pergamon Press, New York (1965)
13. Gutzwiller, M.C.: Chaos in Classical and Quantum Mechanics. Springer, New York (1990)
14. Heller, E.J.: Phys. Rev. Lett. **53**, 1515 (1984)
15. Heller, E.J.: In: Giannoni, M.J., Voros, A., Zinn-Justin, J. (eds.) Chaos and Quantum Physics. Elsevier, Amsterdam (1991)
16. Kaplan, L., Heller, E.J.: Ann. Phys. **264**, 171 (1998)
17. Bogomolny, E.B.: Physica D **31**, 169 (1988)
18. Berry, M.V.: Proc. R. Soc. Lon. A **243**, 219 (1989)
19. Tomsovic, S., Heller, E.J.: Phys. Rev. Lett. **70**, 1405 (1993)
20. Tomsovic, S., Lefebvre, J.H.: Phys. Rev. Lett. **79**, 3629 (1997)
21. Wisniacki, D.A., Vergini, E., Benito, R.M., Borondo, F.: Phys. Rev. E **70**, 035202(R) (2004)
22. Wisniacki, D.A., Vergini, E., Benito, R.M., Borondo, F.: Phys. Rev. Lett. **94**, 054101 (2005)
23. Wisniacki, D.A., Vergini, E., Benito, R.M., Borondo, F.: Phys. Rev. Lett. **97**, 094101 (2006)
24. Keating, J.P., Prado, S.D.: Proc. R. Soc. Lon. A **457**, 1855 (2001)
25. Wisniacki, D., Carlo, G.G.: Phys. Rev. E **77**, 045201(R) (2008)
26. Novaes, M., Pedrosa, J.M., Wisniacki, D., Carlo, G.G., Keating, J.P.: Phys. Rev. E **80**, 035202(R) (2009)
27. Xu, H., Huang, L., Lai, Y.-C., Grebogi, C.: Phys. Rev. Lett. **110**, 064102 (2013)
28. Stöckman, H.J.: Quantum Chaos: An Introduction. Cambridge University Press, Cambridge (2000)
29. Sridhar, S.: Phys. Rev. Lett. **67**, 785 (1991)
30. Stein, J., Stöckmann, H.: Phys. Rev. Lett. **68**, 2867 (1991)
31. Wilkinson, P.B., et al.: Nature (London) **380**, 608 (1996)
32. Akis, R., Ferry, D.K., Bird, J.P.: Phys. Rev. Lett. **79**, 123 (1997)
33. Nöckel, J.U., Stone, A.D.: Nature (London) **385**, 45 (1997)
34. Gmachl, C., Capasso, F., Narimanov, E.E., Nöckel, J.U., Stone, A.D., Faist, J., Sivco, D.L., Cho, A.Y.: Science **280**, 1556 (1998)
35. Lee, S.-B., Lee, J.-H., Chang, J.-S., Moon, H.-J., Kim, S.W., An, K.: Phys. Rev. Lett. **88**, 033903 (2002)
36. Harayama, T., Fukushima, T., Davis, P., Vaccaro, P.O., Miyasaka, T., Nishimura, T., Aida, T.: Phys. Rev. E **67**, 015207(R) (2003)
37. Song, Q.H., Ge, L., Stone, A.D., Cao, H., Wiersig, J., Shim, J.-B., Unterhinninghofen, J., Fang, W., Solomon, G.S.: Phys. Rev. Lett. **105**, 103902 (2010)
38. Wiersig, J., Eberspächer, A., Shim, J.-B., Ryu, J.-W., Shinohara, S., Hentschel, M., Schomerus, H.: Phys. Rev. A **84**, 023845 (2011)
39. Doya, V., Legrand, O., Mortessagne, F., Miniatura, C.: Phys. Rev. Lett. **88**, 014102 (2002)
40. Michel, C., Doya, V., Legrand, O., Mortessagne, F.: Phys. Rev. Lett. **99**, 224101 (2007)
41. Huang, L., Lai, Y.-C., Ferry, D.K., Goodnick, S.M., Akis, R.: Phys. Rev. Lett. **103**, 054101 (2009)
42. Revuelta, F., Benito, R.M., Borondo, F., Vergini, E.G.: Phys. Rev. E **87**, 042921 (2013)
43. Borondo, F., Benito, R.M.: In: Yurtsever, E. (ed.) Frontiers of Chemical Physics. NATO ASI Series C. Kluwer, Dordrecht (1995) (references therein)
44. Esser, R., Tennyson, J., Wormer, P.E.S.: Chem. Phys. Lett. **89**, 223 (1982)
45. Benito, R.M., Borondo, F., Kim, J.-H., Sumpter, B.G., Ezra, G.S.: Chem. Phys. Lett. **161**, 60 (1989)
46. Losada, J.C., Estebaranz, J.M., Benito, R.M., Borondo, F.: J. Chem. Phys. **108**, 63 (1998)
47. Borondo, F., Losada, J.C., Benito, R.M.: Found. Phys. **31**, 147 (2001)
48. Losada, J.C., Benito, R.M., Arranz, F.J., Borondo, F.: Int. J. Quantum Chem. **86**, 167 (2002)
49. Bačić, Z., Light, J.C.: Annu. Rev. Phys. Chem. **40**, 469 (1989)
50. Arranz, F.J., Borondo, F., Benito, R.M.: J. Chem. Phys. **104**, 6401 (1996)
51. Arranz, F.J., Borondo, F., Benito, R.M.: Phys. Rev. Lett. **80**, 944 (1998)
52. Arranz, F.J., Borondo, F., Benito, R.M.: J. Chem. Phys. **107**, 2395 (1997)
53. de Polavieja, G.G., Borondo, F., Benito, R.M.: Phys. Rev. Lett. **73**, 1613 (1994)

54. Sibert III, E.L., Vergini, E., Benito, R.M., Borondo, F.: New J. Phys. **10**, 053016 (2008)
55. Revuelta, F., Vergini, E.G., Benito, R.M., Borondo, F.: Phys. Rev. E **85**, 026214 (1012)
56. Borondo, F., Zembekov, A.A., Benito, R.M.: Chem. Phys. Lett. **246**, 421 (1995)
57. Borondo, F., Zembekov, A.A., Benito, R.M.: J. Chem. Phys. **106**, 5068 (1996)
58. Borondo, F., Zembekov, A.A., Benito, R.M.: J. Chem. Phys. **107**, 7934 (1997)
59. Vergini, E.G., Sibert III, E.L., Revuelta, F., Benito, R.M., Borondo, F.: Europhys. Lett. **89**, 40013 (2010)

Chaos Powered Symbolic Regression in Be Stars Spectra Modeling

Ivan Zelinka, Lenka Skanderova, Petr Saloun, Roman Senkerik and Michal Pluhacek

Abstract Be stars are characterized by prominent emission lines in their spectrum. In the past research has attention been given to creation a feature extraction method for classification of Be stars with focusing on the automated classification of Be stars based on typical shapes of their emission lines. The aim was to design a reduced, specific set of features characterizing and discriminating the shapes of Be lines. In this chapter we discuss possibility to create in an evolutionary way the model of spectra of Be stars. We focus on the evolutionary synthesis of the mathematical models of Be stars based on typical shapes of their emission lines. Analytical programming powered by classical random as well as chaotic random-like number generator is used here. Experimental data are used from the archive of the Astronomical Institute of the Academy of Sciences of the Czech Republic. Interpretation and explanation of analysis is given and discussed in this chapter.

1 Introduction

Technological progress and growing computing power are causing data avalanche in almost all sciences, including astronomy. The full exploitation of these massive distributed data sets clearly requires automated methods. One of the difficulties is the inherent size and dimensionality of the data. The efficient classification requires that we reduce the dimensionality of the data in a way that preserves as many of the physical correlations as possible.

I. Zelinka (✉) · L. Skanderova · P. Saloun
VSB-Technical University of Ostrava, 17.listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
e-mail: ivan.zelinka@vsb.cz

R. Senkerik · M. Pluhacek
Tomas Bata University in Zlin, Zlin, Czech Republic
e-mail: senkerik@fai.utb.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_13,

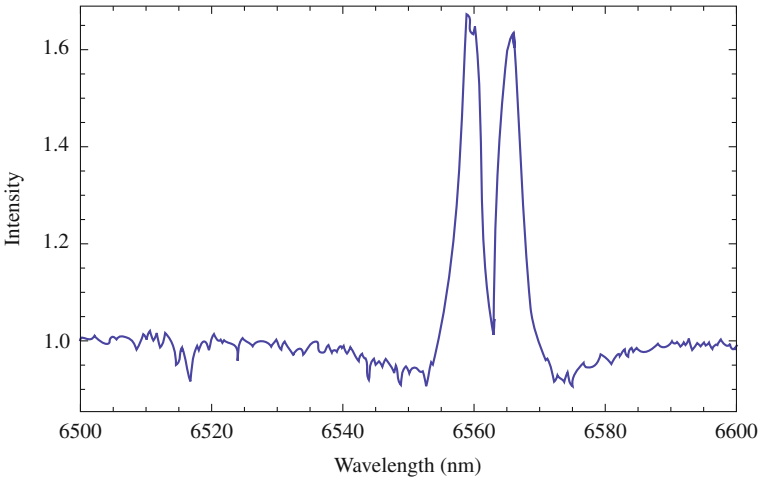


Fig. 1 Typical shapes of emission lines in spectra of Be stars

Be stars are hot, rapidly rotating B-type stars with equatorial gaseous disk producing prominent emission lines H_α in their spectrum [1]. Be stars show a number of different shapes of the emission lines, as we can see in Fig. 1. These variations reflect underlying physical properties of a star.

Our work is focused on the evolutionary-automated model synthesis of Be stars based on typical shapes of their emission lines. There has not been much work on classification of Be stars. The only application found [2] is focused on a broader category of variable stars including pulsating Be stars. However, the method is not suitable for our goals, as it is applied on the whole spectrum where the local differences in the shapes of Be lines are lost. Another approach is in [3] where the to zoom at the small part of a spectrum with the Be line is applied, specific set of features characterizing and discriminating the shapes of Be lines. Data of Be stars spectra come from the archive of the Astronomical Institute of the Academy of Sciences of the Czech Republic.

2 Used Methods

For our experiments described here standard hardware and algorithms has been used. All important information about algorithms used in our experiments are mentioned and referred here.

Table 1 Algorithms setting

DE		SOMA	
NP	500	PopSize	500
Dimensions	100	Dimensions	100
Generations	500	Migrations	10
F	0.9	PRT	0.1
CR	0.5	PathLength	5
		Step	0.21

2.1 Evolutionary Algorithms

Comparing to the previous method that is based on the wavelet transform and its power spectrum, in this research is used symbolic (i this case AP) regression with evolutionary algorithms like Self-Organising Migrating Algorithm (SOMA) and Differential Evolution (DE). While in the [3] a resulting feature vector is composed of two parts: 1. wavelet power spectrum, 2. value indicating the orientation of the H_α line (this information is lost in the wavelet power spectrum), we are focused on evolutionary model synthesis that shall create model of observed spectra. Together with SOMA and DE has been used approach called Analytic Programming (AP) [4]. AP has been powered by chaotic number generator and compared with the same AP powered by classical pseudorandom number generator. Used algorithms were set according to (Table 1).

2.2 Symbolic Regression

In [4] is discussed an alternative approach for symbolic structures and solutions synthesis, used here as well as brief well known methods. For example Genetic Programming (GP), [5, 6] or Grammatical Evolution (GE), [7, 8]. Generally, there are two well known methods, which can be used for symbolic structures synthesis by means of computers. The first one is called GP and the other is GE. Another interesting research was carried out by Artificial Immune Systems (AIS) or/and systems, which do not use tree structures like linear GP and other similar algorithm like Multi Expression Programming (MEP), etc. In this chapter, a different method called Analytic Programming (AP), is presented. AP is a grammar free algorithmic superstructure, which can be used by any programming language and also by any arbitrary Evolutionary Algorithm (EA) or another class of numerical optimization method. This chapter describes use and results of AP use with EA's like Differential Evolution (DE), Self-Organising Migrating Algorithm (SOMA) on Be-Stars spectra model synthesis. All case studies has been carefully prepared and repeated in order to get valid statistical data for proper conclusions (Table 2).

The initial idea of symbolic regression by means of a computer program was proposed in GP [5, 6]. The other approach of GE was developed in [8] and AP in

Table 2 Minimal, maximal and average fitness of synthesized models

	DE PRNGs	DE Chaos	SOMA PRNGs	SOMA Chaos
Max	0.960	0.783	0.692	0.431
Avg	0.352	0.238	0.231	0.132
Min	0.253	0.143	0.195	0.096

[9]. Another interesting investigation using symbolic regression were carried out in [10] on AIS and Probabilistic Incremental Program Evolution (PIPE), which generates functional programs from an adaptive probability distribution over all possible programs. Yet another new technique is the so called *Transplant Evolution*, see [11–13] which is closely associated with the conceptual paradigm of AP, and modified for GE. GE was also extended to include DE by [14]. Generally speaking, it is a process which combines, evaluates and creates more complex structures based on some elementary and noncomplex objects, in an evolutionary way. Such elementary objects are usually simple mathematical operators $(+, -, \times, \ldots)$, simple functions $(sin, cos, And, Not, \ldots)$, user-defined functions (simple commands for robots—MoveLeft, TurnRight, ...), etc. An output of symbolic regression is a more complex "object" (formula, function, command, ...), solving a given problem like data fitting of the so-called Sextic and Quintic problem [15, 16], randomly synthesized function [16], Boolean problems of parity and symmetry solution (basically logical circuits synthesis) [9, 17], or synthesis of quite complex robot control command by [6, 18]. Examples mentioned in [4] are just a few samples from numerous repeated experiments done by AP, which are used to demonstrate how complex structures can be produced by symbolic regression in general for different problems, see [4].

2.3 Experiment Design

Our experiments has been set so that analytic programming powered by classical pseudorandom number generator and deterministic chaos generators (see for example [19–21]), were used. Based on the fact that numerical precision has impact on existence of periodicity in deterministic chaos [19–21], we have selected logistic Eq. 1, and data series generated by this equation with setting $A = 4$. Algorithms selected for our experiments were SOMA [22] and differential evolution (DER and 1Bin) [23].

$$x_{n+1} = Ax_n\,(1 - x_n) \tag{1}$$

The cost function has been defined according to Eq. 2 and the main aim of the used evolution was to find formula, that gives the smallest value of Eq. 2.

Table 3 Periodicity dependance of Eq. 1 on various precision

Precision	Maximal period
1	4
2	10
3	29
4	36
5	170
6	481
7	758
8	4514
9	11227
10	35200
11	57639
12	489154
13	518694

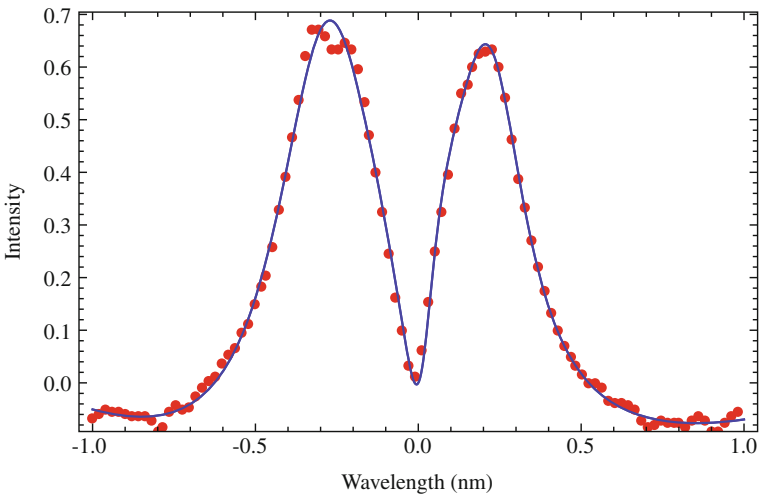


Fig. 2 Typical result of evolutionary fitting of shapes of emission lines in spectra of Be stars, fitted by formula (3)

$$\sum_{1}^{n} |data_i - synthesized\_data_i| \tag{2}$$

All experiments were done in Mathematica 9, on MacBook Pro, 2.8 GHz Intel Core 2 Duo. Test data of the Be stars spectra comes from the archive of the Astronomical Institute of the Academy of Sciences of the Czech Republic. Because this is pioneering experiment, we have used one data set and repeat each experiment 100 times. The aim was to find suitable model of Be star spectra and mainly check whether more intensive experiments shall be done in this way. In total, 400

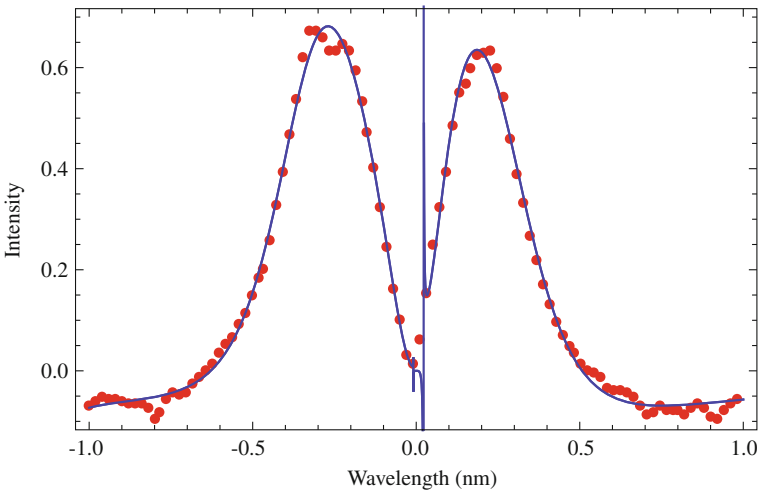


Fig. 3 Another result of evolutionary fitting of shapes of emission lines in spectra of Be stars, fitted by formula (4)

(2 algorithms × 100 repetitions × 2 different pseudorandom number generators—PRNGs) evolutionary experiments has been done. In each experiment, the PRNGs used only on the start of Eq. 1 to set initial condition x_{start}. Remaining use of Eq. 1 was PRNGs free, i.e. PRNGs was not further in use.

3 Results

The results of our experiments, based on AP with SOMA and DE are summarized in the Table 2 and visualized on Figs. 2 and 3.

$$-x^2\left(-\frac{16.485-3.12223x}{(73.8683x^2-x)\left(2.24149(x-0.962468)+1.12563x-\frac{x-0.407234}{x+0.862429}+2.17075\right)+0.159111}\right.$$
$$\left.-\frac{x\left(-0.277185+x-\frac{0.0233139}{x}\right)-0.894691(x+1.84093)x+2x-0.0228628}{-\frac{x^2}{1.52912-x}+x(1.47865-1.11216x)+(2x-0.519239)x-x}\right) \tag{3}$$

$$\frac{0.398423(0.516715x+1.1429)x\times\left(x+\left(-\frac{0.939778(x-2.35703)}{x}-\frac{3.85544}{x}\right)(4.48681x-2.15497)\right)}{(-19.2933(2x+1.56706)x^2+(x-6.35439)\left(-x^2-3x\right)+x-6.53511)} \tag{4}$$
$$\times\left(-\frac{\left(\frac{\frac{1.04589}{1.04625-x}-0.982957}{x}-0.908173\right)\left(\frac{1.06601(x+1.30534)-x^2}{2.19415+\frac{0.000317128}{x}}-0.640813\right)}{x(2.85538(x^2+0.278391x)+x)}-x\right)$$

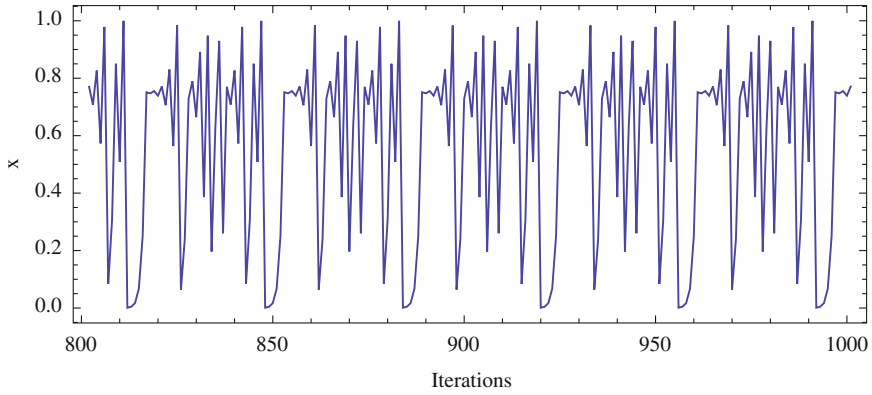


Fig. 4 Time series of period 36 (precision = 4) based on Eq. 1 for $A = 4$, see Table 3

4 Conclusion

The main goal of this research was to test whether it is possible to use AP on Be-stars spectra modeling. Two version of AP has been used. The first one was powered by classical pseudorandom number generator and the second one by pseudo-chaotic number generator, in this case by logistic equation. Ap was used with SOMA as well as with DE algorithms. Obtained results (see Table 2 and Figs. 2 and 3) shows that AP can be used for that kind of task and both cases, i.e. AP with PRNGS as well as AP with chaos can be used.

Based on results from [19–21], see also Table 3, we are going to use chaotic generators with variable level of chaos in order to get more results that show dependance of AP performance on numerical precision of chaotic generators. As was reported in [19–21], it is possible to use deterministic chaos generators instead of pseudo-random number and set by control parameter periodicity of obtained data series. It is reported in [19–21], see also Table 3 and Fig. 4. More information about mutual fusion of evolutionary algorithms and deterministic chaos can be found in [24, 25], while about AP in [4].

Acknowledgments The following two grants are acknowledged for the financial support provided for this research: Grant Agency of the Czech Republic—GACR P103/13/08195S, by the Development of human resources in research and development of latest soft computing methods and their application in practice project, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic, partially supported by Grant of SGS No. SP2013/114, VŠB—Technical University of Ostrava, Czech Republic, and by European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089.

References

1. Thizy, O.: Classical Be Stars High Resolution Spectroscopy, Society for Astronomical Sciences Annual Symposium, pp. 27–49. http://adsabs.harvard.edu/abs/2008SASS%8527%8549T, Provided by the SAO/NASA Astrophysics Data System (2008)
2. Debosscher, J.: Automated Classification of Variable Stars: Application to the OGLE and CoRoT Databases. Institute of Astronomy, Faculty of Sciences, Catholic University of Leuven, Leuven (2009)
3. Bromova, P., Skoda, P., Zendulka, J.: Wavelet based feature extraction for clustering of Be stars. In: Proceedings of Nostradamus 2013: International Conference Prediction, Modeling and Analysis of Complex Systems, Springer Series: Advances in Intelligent Systems and Computing, vol. 210, pp. 467–474 (2013)
4. Zelinka, I., Davendra, D., Senkerik, R., Jasek, R., Oplatkova, Z: (2011). Analytical programming—a novel approach for evolutionary synthesis of symbolic structures. In: Kita, E. (ed.) Evolutionary Algorithms. ISBN: 978-953-307-171-8, InTech, doi:10.5772/16166. http://www.intechopen.com/books/evolutionary-algorithms/analytical-programming-a-novel-approach-for-evolutionary-synthesis-of-symbolic-structures
5. Koza, J.: Genetic programming: a paradigm for genetically breeding populations of computer programs to solve problems. Stanford University, Computer Science Department, Technical Report, STAN-CS-90-1314 (1990)
6. Koza, J.: Genetic Programming. MIT Press, Cambridge (1998)
7. O'Neill, M., Ryan, C.: Grammatical Evolution, Evolutionary Automatic Programming in an Arbitrary Language. Springer, New York (2003)
8. Ryan, C., Collins, J., O'Neill, M.: Grammatical Evolution: Evolving Programs for an Arbitrary Language. Lecture Notes in Computer Science, First European Workshop on Genetic Programming (1998)
9. Zelinka, I., Oplatkova, Z., Nolle, L.: Analytic programming—symbolic regression by means of arbitrary evolutionary algorithms. Int. J. Simul. Syst. Sci. Technol. **6**(9), 44–56 (2005)
10. Johnson, C.: Artificial immune systems programming for symbolic regression. In: Ryan, C., Soule, T., Keijzer, M., Tsang, E., Poliand, R., Costa, E. (eds.) Lecture Notes in Computer Science, pp. 345–353. Springer, Berlin (2004)
11. Weisser, R., Osmera, P.: Two-level transplant evolution for optimization of general controllers. In: New Trends in Technologies. Sciyo, Croatia (2010)
12. Weisser, R., Osmera, P.: Two-level tranpslant evolution. In: 17th Zittau Fuzzy Colloquium, Zittau, Germany (2010)
13. Weisser, R., Osmera, P., Matousek, R.: Transplant evolution with modified schema of differential evolution: optimization structure of controllers. In: International Conference on Soft Computing MENDEL, Brno, Czech Republic (2010)
14. O'Neill, M., Brabazon, A.: Grammatical differential evolution. In: Proceedings of International Conference on Artificial Intelligence, pp. 231–236. CSEA Press (2006)
15. Koza, J., Bennet, F., Andre, D., Keane, M.: Genetic Programming III. Morgan Kaufmann, New York (1999)
16. Zelinka, I., Oplatkova, Z.: Analytic programming—comparative study. In: Proceedings of Second International Conference on Computational Intelligence, Robotics, and Autonomous Systems, Singapore (2003)
17. Koza, J., Keane, M., Streeter, M.: Evolving inventions. Sci. Am. 40–47 (2003)
18. Oplatkova, Z., Zelinka, I.: Investigation on artificial ant using analytic programming. In: Proceedings of Genetic and Evolutionary Computation Conference, pp. 949–950, Seattle, WA (2006)
19. Zelinka, I., Senkerik, R., Pluhacek, M.: Do evolutionary algorithms indeed require randomness? In: IEEE Congress on Evolutionary Computation, pp. 2283–2289. Cancun, Mexico (2013)

20. Zelinka, I., Chadli, M., Davendra, D., Senkerik, R., Pluhacek, M., Lampinen, J.: Hidden periodicity—chaos dependance on numerical precision. In: Proceedings of Nostradamus 2013: International Conference Prediction, Modeling and Analysis of Complex Systems, Springer Series: Advances in Intelligent Systems and Computing, vol. 210, pp. 47–59 (2013)
21. Zelinka, I., Chadli, M., Davendra, D., Senkerik, R., Pluhacek, M., Lampinen, J.: Do evolutionary algorithms indeed require random numbers? extended study. In: Proceedings of Nostradamus 2013: International Conference Prediction, Modeling and Analysis of Complex Systems, Springer Series: Advances in Intelligent Systems and Computing, vol. 210, pp. 61–75 (2013)
22. Zelinka I.: SOMA—self organizing migrating algorithm. In: Babu, B.V., Onwubolu, G., (eds.) New Optimization Techniques in Engineering, pp. 167–218. Springer, New York (2004)
23. Price K.: An introduction to differential evolution. In: Corne, D., Dorigo, M., Glover, F., (eds.) New Ideas in Optimization, pp. 79–108. McGraw-Hill, London (1999)
24. Zelinka, I., Chen, G., Celikovsky, S.: Chaos Synthesis by means of evolutionary algorithms. Int. J. Bifurcat. Chaos **18**(4), 911–942 (2008). ISSN 0218–1274
25. Zelinka, I., Chen, G., Celikovsky, S.: Evolutionary Algorithms and Chaotic Systems. Springer, Germany (2010)

Mathematical Modeling of Heat Loss of a Sphere in Contact with a Well Stirred Fluid

Juan Carlos Beltrán-Prieto and Karel Kolomazník

Abstract The knowledge of several transport properties is important especially when heat transfer has to be evaluated. However, not always the data such as thermal diffusivity is known or available for specific types of material. In this paper, a study was performed to describe the equations that lead to the calculus of thermal diffusivity, heat transfer and calculus of temperature in function of time of a sphere particle that is immersed in a well stirred fluid. The model obtained was tested by comparing the predicted thermal diffusivity value of an orange and the real value.

Keywords Thermal diffusivity · Heat transfer · Stirred fluid · Conductivity · Spherical particle

1 Introduction

This paper describes the modeling of conductive cooling of a sphere that is immersed in a well stirred media to estimate the thermal diffusivity of the solid, the heat transfer and the temperature of the fluid at a specific time. The study is aimed to provide a fundamental understanding of this cooling process that can be of assistance when it is desired to control parameters, improve the functionality and operation of cooling, or for performance prediction. Broad applications are found, particularly at micro and nanoscale level, where the cooling of microspheres, microchannels, microcylinders or nanoparticles has recently gained attention. Additionally, conductive cooling is of extreme importance as it is the main mechanism involved in the use of

J. C. Beltrán-Prieto (✉) · K. Kolomazník
Department of Automation and Control Engineering, Faculty of Applied Informatics, Tomas Bata University in Zlín, Nám. T. G. Masaryka 5555, 760 01 Zlín, Czech Republic
e-mail: prieto@fai.utb.cz

K. Kolomazník
e-mail: kolomaznik@fai.utb.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_14,

laser-induced incandescence technique for the determination of the size of Fe or MnO_2 nanoparticles [1] or in the characterization of the diameter of particles such as carbon black [2]. The complete understanding of the controlling the heat transfer in a system lies in finding effective ways to model the cooling process. In order to perform the mathematical model, it was analyzed that a homogeneous solid sphere of radius (R), with an initial uniform temperature (T_1,) was suddenly immersed at time $t=0$ in a volume (V_f) of a well stirred fluid of temperature (T_0) in an insulated tank. The procedure described here aims to find the thermal diffusivity ($\alpha_s = k_s/ \rho_s c_{p,s}$) of the solid by observing the change of the fluid temperature T_f with time. The parameter k_s represents the thermal conductivity, ρ_s the density and $c_{p,s}$ the specific heat capacity of the solid.

2 Methodology

The approach for the modeling started from the consideration of the energy equation and Fourier's law of heat conduction for solids. The solid was regarded to be a spherical particle. After introducing the dimensional variables in the equations, boundary conditions at the center of the sphere and at the surface were taken into account. After, an energy balance for the fluid in the tank and an energy total balance for the sphere in the tank were realized. In order to calculate the variation of energy in the fluid, dimensional equations were obtained by substitution of the corresponding dimensional variables (fluid temperature, distance and time). This new equations were solved by applying the Laplace transform solution method, followed by the application of Bessel function. The increment in the fluid temperature allows the determination of the thermal diffusivity of the sphere. In order to know the heat transfer, the derivative of the equation that described the temperature in the fluid was derived. The model was tested by analyzing mathematically the cooling of an orange immersed in a well stirred fluid.

3 Description of the Modeling

The analysis starts by considering the energy equation combined with Fourier's law of heat conduction for solids, as described in (1). Equation (2) arises from stating that thermal conductivity is independent of the temperature and position. By considering spherical coordinates, Eq. (3) is obtained

$$\rho_s, c_{p,s} \frac{\partial T}{\partial t} = (\nabla \cdot k_s \nabla T) \tag{1}$$

$$\frac{\partial T}{\partial t} = (\alpha \nabla^2 T) \tag{2}$$

$$\rho_s \cdot c_{p,s} \left(\frac{\partial T}{\partial t} \right) = k_s \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial T}{\partial r} \right) \right] \tag{3}$$

Introducing the dimensional variables or respective derivatives of temperature of sphere (4), dimension (5), time (6), and thermal diffusivity (7) and simplifying, the equation (3) takes the form of (8):

$$\Theta_s = \frac{T_1 - T_s}{T_1 - T_0} \tag{4}$$

$$\xi = \frac{r}{R} \tag{5}$$

$$\tau = \frac{\alpha_s t}{R^2} \tag{6}$$

$$\alpha = \frac{k_s}{\rho cp} \tag{7}$$

$$\left(\frac{\partial \Theta_s}{\partial \tau} \right) = \left[\frac{1}{\xi^2} \frac{\partial}{\partial \xi} \left(\xi^2 \frac{\partial \Theta_s}{\partial \xi} \right) \right] \tag{8}$$

Due to the fact that it is a second order balance with respect to the boundary spatial variable, two boundary conditions are necessary, one is at the center of the sphere (9), where symmetry conditions must be accomplished, and the other is at the particle exterior surface (10), where the sphere surface and the fluid have the same temperature if perfect mixing exists. Normally one condition for time is required ($t = 0$) and two for the fixed spatial position ($r = 0$ and $r = R$). At time zero ($t = 0$), the temperature of the sphere (T_s) corresponds to the initial temperature (T_1), ($T_s = T_1$)

$$r = 0 \quad \left. \frac{dT_s}{dr} \right|_{r=0} = 0 \rightarrow\rightarrow \xi = 0 \quad \left. \frac{\partial \Theta_s}{\partial \xi} \right|_{r=0} \tag{9}$$

$$r = R \quad T_{solid} = T_{fluid} \rightarrow\rightarrow \xi = 1 \quad \Theta_s = \Theta_f \tag{10}$$

The temperature in the sphere changes according to the time because the tank is finite; therefore, in order to know the variation of temperature of the sphere in function of the time it is necessary to realize an energy balance for the fluid in the tank. Introducing the term V/V to include the mass (ms) and density of the sphere, we get (12). Q represents the heat transfer from the sphere to the fluid

$$Q = \left(4\pi R^2\right)\left(k_s \frac{\partial T_s}{\partial r}\right) \tag{11}$$

$$\left[\left(4\pi R^2\right)\left(k_s \frac{\partial T_s}{\partial r}\right)\right]\left[\frac{\frac{m_s}{\rho_s}}{\frac{4}{3}\pi R^3}\right] = \frac{3}{R}\frac{m_s}{\rho_s}k_s \frac{\partial T_s}{\partial r} \tag{12}$$

Knowing the loss of energy, the energy total balance for the sphere in the tank is (13):

$$V_{fluid}\frac{dT_{fluid}}{dT} = -\frac{3}{R}\frac{m_s}{\rho_s}k_s \frac{\partial T_s}{\partial r} \tag{13}$$

where V_{fluid} is the tank volume excluding the sphere. To be consistent with the units it is necessary to add density (ρ_{fluid}) and specific heat ($c_{p,fluid}$) of fluid. After, it would be useful to substitute the volume and express k_s in terms of density, specific heat and thermal diffusivity of sphere as described in (14)

$$(V_{fluid})\left(\frac{dT_{fluid}}{dt}\right)\left(\rho_{fluid} \cdot c_{p,fluid}\right) = -\frac{3}{R}V_s \alpha \rho_s cp_s \frac{\partial T_s}{\partial r} \tag{14}$$

In order to find the variation of energy in the fluid, it is necessary to substitute the corresponding dimensional variables of fluid temperature (T_{fluid}) (15), distance (5) and time (6), in Eq. (14) to get (16). For this last equation, the initial condition is that at $\tau = 0$, $\Theta_F = 1$

$$T_{fluid} = T_1 - \Theta_{fluid}\left(T_1 - T_0\right) \tag{15}$$

$$\left(\frac{\partial \Theta_{fluid}}{\partial \tau}\right) = \frac{-3V_s \rho_s cp_s}{\left(V_{fluid}\right)\left(\rho_{fluid} \cdot cp_{fluid}\right)}\frac{\partial \Theta_s}{\partial \xi} \tag{16}$$

The groups of Eqs. (8) and (16) can be solved by the Laplace transform solution method. For the solution of (8), we need to consider $\Theta_s = u$ and $\tau = 0$. The solution is presented in Eq. (17).

$$\left\{\left(\frac{d^2 u}{d\xi^2}\right) + \frac{2}{\xi}\frac{du}{d\xi}\right\} - Su = 0 \tag{17}$$

Equation (17) has to be multiplied by ξ^2 (22) to be solved by Bessel function as presented in Eq. (18):

$$u = \left(\sqrt{\frac{2}{\pi\sqrt{s}}}\right)\left(\frac{1}{\xi}\right)\left(C_1 senh\sqrt{s}\xi\right) \tag{18}$$

$$\frac{du}{d\xi} = \left(\sqrt{\frac{2}{\pi\sqrt{s}}}\right) C_1 \left[\left(\frac{1}{\xi}\right)\sqrt{s}\left(cosh\sqrt{s}\xi\right) - senh\sqrt{s}\xi\left(\frac{1}{\xi^2}\right)\right] = \frac{\mathrm{d}\Theta_\mathrm{s}}{d\xi} \tag{19}$$

For the solution of (16), we consider $\Theta_{\mathrm{fluid}} = \mathrm{u}_{\mathrm{fluid}}$ and $\tau = \mathrm{s}$, which leads to Eqs. (20) and (21).

$$S \cdot \mathrm{u}_{\mathrm{fluid}} - (\Theta_{\mathrm{fluid}} \mid \tau = 0 \mid) = \frac{-3V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})} \frac{\partial\Theta_\mathrm{s}}{\partial\xi} \tag{20}$$

$$S \cdot \mathrm{u}_{\mathrm{fluid}} - 1 = \frac{-3V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})} \frac{\partial\Theta_\mathrm{s}}{\partial\xi} \tag{21}$$

Substituting (19) in (21) we get (22) to know $\mathrm{u}_{\mathrm{fluid}}$:

$$\mathrm{u}_{\mathrm{fluid}} = \left\{\left(\frac{1}{\mathrm{s}}\right)\left(\frac{-3V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}\right)\left(\sqrt{\frac{2}{\pi\sqrt{s}}}\right) C_1\left[\left(\frac{1}{\xi}\right)\sqrt{s}\left(cosh\sqrt{s}\xi\right) - senh\sqrt{s}\xi\left(\frac{1}{\xi^2}\right)\right]\right\} + \frac{1}{\mathrm{s}} \tag{22}$$

For $\xi = 1$:

$$\mathrm{u}_{\mathrm{fluid}} = \left\{\left(\frac{1}{\mathrm{s}}\right)\left(\frac{-3V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}\right)\left(\sqrt{\frac{2}{\pi\sqrt{s}}}\right) C_1\left[\left(\frac{1}{\xi}\right)\sqrt{s}\left(cosh\sqrt{s}\right) - senh\sqrt{s}\right]\right\} + \frac{1}{\mathrm{s}} \tag{23}$$

Equations (18) and (23) contains the term C1. To find the corresponding value, we apply the second boundary condition in the sphere ($\xi = 1$ and $\Theta_\mathrm{s} = \Theta_{\mathrm{fluid}}$). After, by solving simultaneously both equations, simplifying and putting in terms of, we obtain (24)

$$\mathrm{u}_{\mathrm{fluid}} = \left[\left(\frac{-3}{\mathrm{s}}\right)\left(\frac{V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}\right) \frac{[\sqrt{s} - tanh\sqrt{s}]}{\left((s)tanh\sqrt{s} + \left(\frac{3V_s\rho_s cp_s}{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}\right)[\sqrt{s} - tanh\sqrt{s}]\right)}\right] + \frac{1}{\mathrm{s}} \tag{24}$$

Equation (24) can be transformed into (25):

$$u_{fluid} = -\frac{3}{\sqrt{s}}\left[\frac{\left[\sqrt{s} - tanh\sqrt{s}\right]}{\left(\left[\frac{(V_{fluid})(\rho_{fluid}\cdot cp_{fluid})}{V_s\rho_s cp_s}\right]s\sqrt{s}tanh\sqrt{s} + (3)\left[s - \sqrt{s}\cdot tanh\sqrt{s}\right]\right)}\right] + \frac{1}{s} \quad (25)$$

Changing the signs and factorizing $\sqrt{s}\cdot tanh\sqrt{s}$ from the denominator, and dividing the numerator by $\sqrt{s}$, Eq. (25) becomes:

$$u_{fluid} = 3\left[\frac{\left[1 - \frac{tanh\sqrt{s}}{\sqrt{s}}\right]}{\left(\sqrt{s}tanh\sqrt{s}\left\{\left[\frac{-(V_{fluid})(\rho_{fluid}\cdot cp_{fluid})}{V_s\rho_s cp_s}\right]s + 3\right\} - (3)[s]\right)}\right] + \frac{1}{s} \quad (26)$$

To solve this equation it is necessary to apply the Inverse Laplace transformation ($\mathcal{L}^{-1}$) to both terms of the equation.

$$u_{fluid} = 3\mathcal{L}^{-1}\left[\frac{\left[1 - \frac{tanhh\sqrt{s}}{\sqrt{s}}\right]}{\left(\sqrt{s}tanh\sqrt{s}\left\{\left[\frac{-(V_{fluid})(\rho_{fluid}\cdot cp_{fluid})}{V_s\rho_s cp_s}\right]s + 3\right\} - (3)[s]\right)}\right] + \mathcal{L}^{-1}\left(\frac{1}{s}\right) \quad (27)$$

$$u_{fluid} = 3\mathcal{L}\left[\frac{\left[1 - \frac{tanh\sqrt{s}}{\sqrt{s}}\right]}{\left(\sqrt{s}tanh\sqrt{s}\left\{\left[\frac{-(V_{fluid})(\rho_{fluid}\cdot cp_{fluid})}{V_s\rho_s cp_s}\right]s + 3\right\} - (3)[s]\right)}\right] + 1 \quad (28)$$

It is possible to transform the previous equation by using the Heaviside theorem of partial fraction expansion. Firstly, the numerator and denominator of the first term are divided by (1/s) leading to (29). This equation can be represented as shown in (30).

$$u_{fluid} = 3\mathcal{L}^{-1}\left[\frac{\left[\frac{1}{s} - \frac{tanh\sqrt{s}}{s^{3/2}}\right]}{\left(\frac{tanh\sqrt{s}}{\sqrt{s}}\left\{\left[\frac{-(V_{fluid})(\rho_{fluid}\cdot cp_{fluid})}{V_s\rho_s cp_s}\right]s + 3\right\} - (3)\right)}\right] + 1 \quad (29)$$

$$u_{fluid} = 3\mathcal{L}^{-1}\left[\frac{N(s)}{D(s)}\right] + 1 \quad (30)$$

The denominator contains one root for s = 0 and roots for $\sqrt{s_k} = i\mu_k$, for $1 \leq k \leq \infty$, and s_k represents the roots of $\tan s_k = \frac{3\mu_k}{3+\left(\frac{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}{V_s \rho_s cp_s}\right)\mu_k^2}$.

Therefore $\left[\frac{N(0)}{D'(0)}\right] = -\frac{B/3}{1+B}$ and $\left[\frac{N(s_k)}{D'(s_k)}\right] = \frac{2B}{9(1+B)+B^2\mu_{k^2}}$ are obtained from where $-B = \frac{(V_{fluid})(\rho_{fluid} \cdot cp_{fluid})}{V_s \rho_s cp_s}$.

To continue with the solution of Eq. (29), the Heaviside partial fraction expansion theorem is applied to get (31):

$$u_{fluid} = \left[\frac{B}{1+B} + 6B\sum_{k=1}^{\infty} \frac{e^{\mu_{k^2 s}}}{9(1+B)+B^2\mu_{k^2}}\right] + 1 \tag{31}$$

Turning back to the variables Θ_{fluid} and τ established previously, Eq. (32) is obtained. After, we express the equation in terms of time and temperatures as expressed in (33)

$$\Theta_{fluid} = \left[\frac{B}{1+B} + 6B\sum_{k=1}^{\infty} \frac{e^{\mu_{k^2 \tau}}}{9(1+B)+B^2\mu_{k^2}}\right] + 1 \tag{32}$$

$$\frac{T_1 - T_{fluid}}{T_1 - T_0} = \left[\frac{B}{1+B} + 6B\sum_{k=1}^{\infty} \frac{e^{\mu_{k^2}\frac{\alpha_s t}{R^2}}}{9(1+B)+B^2\mu_{k^2}}\right] + 1 \tag{33}$$

The increment in the fluid temperature allows the determination of the thermal diffusivity of the sphere. Equation (34) is obtained to determine the temperature in the fluid in function of time.

$$T_{fluid} = T_1 + \{T_0 - T_1\}\left\{\left[\frac{B}{1+B} + 6B\sum_{k=1}^{\infty} \frac{e^{\mu_k^2\frac{\alpha_s t}{R^2}}}{9(1+B)+B^2\mu_{k^2}}\right] + 1\right\} \tag{34}$$

To calculate the heat transfer (35) we need to calculate the derivative of (36)

$$q_r = -k\frac{\partial T}{\partial r}|_{r=R}(4\pi R^2) \tag{35}$$

$$\frac{dT}{dr} = [T_0 - T_1]\left[6B\sum_{k=1}^{\infty} \frac{-2t\alpha_s\mu_{k^2}}{R^3}\frac{e^{\mu_{k^2}\frac{\alpha_s t}{R^2}}}{9(1+B)+B^2\mu_{k^2}}\right] \tag{36}$$

$$q_r = -k\left(4\pi R^2\right)[T_0 - T_1]\left[6B\sum_{k=1}^{\infty} \frac{-2t\alpha_s\mu_{k^2}}{R^3}\frac{e^{\mu_{k^2}\frac{\alpha_s t}{R^2}}}{9(1+B)+B^2\mu_{k^2}}\right] \tag{37}$$

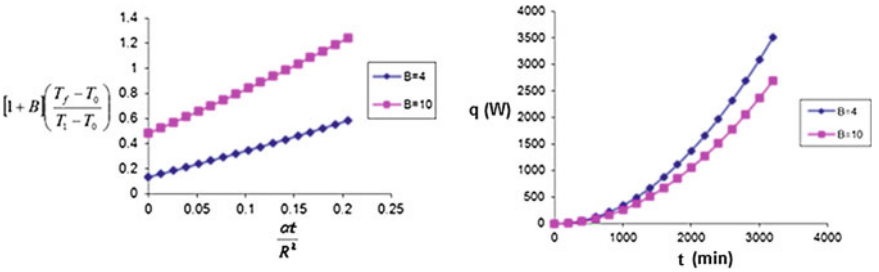


Fig. 1 Variation of temperature of an orange when is immersed in a well stirred fluid (**a**) and effect of B values on heat transfer (**b**)

3.1 Application

Equations (34) and (37) were used to study the cooling of an orange immersed in a well stirred fluid. Chemical properties of orange and water were obtained from [3] and [4]:

$$R = 0.045\,\text{m},\ T_0 = 25\,^{\circ}\text{C},\ \rho_{orange} = 950\,\text{kg/m}^3,\ h_{orange} = 15\,\text{W/m}^2\cdot\text{K},$$
$$k_{orange} = 0.431\,\text{W/m}^2\cdot\text{K},\ C_{p,orange} = 3.77\,\text{kJ/kg}\cdot\text{K},\ \alpha = 1.3\times10^{-7}\,\text{m}^2/\text{s},$$
$$K_{H_2O} = 0.59\,\text{W/m}^2\cdot\text{K},\ C_{p,H_2O} = 4.189\,\text{kJ/kg}\cdot\text{K}$$

In this analysis two values of B are considering (B = 4 and B = 10), higher value of B is an indicative, for example, of bigger tank dimensions. Figure 1a presents the variation of temperature of the orange of radius R and at temperature T_1 when it is immersed in a well stirred fluid of temperature T_0, in relation to the time. Figure 1b shows that a greater heat interchange is achieved with a lower value of B.

It is possible to use Eq. (34) to estimate the thermal diffusivity of the orange, in this case, it is proposed to find it by using the known values of its properties and compare the T_f values with the estimation. By doing this process, the diffusivity value found was $1.17\times10^{-7}\,\text{m}^2/\text{s}$, a closer value to data reported ($1.3\times10^{-7}\,\text{m}^2/\text{s}$) [4]

4 Conclusion

The series of equations that lead to the modeling of loss of heat of a spherical particle when it is immersed in a well stirred fluid was performed. Equations for thermal diffusivity, heat transfer and for the calculus of temperature of the sphere in function of time were obtained. Concordance between the value of thermal diffusivity predicted by the model and data reported for an orange was observed.

Acknowledgments The project was financially supported by the European Regional Development Fund under the Project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089 and Tomas Bata University

Internal Grant (IGA/FAI/2013/037/) The author is also grateful for the doctoral scholarship provided by the National Council of Science and Technology (CONACYT) in Mexico and Tomas Bata University in Zlín.

References

1. Lehre, T., Suntz, R., Bockhorn, H.: Time-resolved two-color LII: size distributions of nanoparticles from gas-to-particle synthesis. Proc. Combust. Inst. **30**, 2585–2593 (2005)
2. Liu, F., Stagg, B.J., Snelling, D.R., Smallwood, G.J.: Effects of primary soot particle size distribution on the temperature of soot particles heated by a nanosecond pulsed laser in an atmospheric laminar diffusion flame. Int. J. Heat Mass Transfer. **49**, 777–788 (2006)
3. Geankoplis, C.J.: Transport Processes and Separation Process Principles, 4th edn. Prentice Hall, New Jersey (2003)
4. Cengel, Y.A.: Heat Transfer: A Practical Approach, 2nd edn. McGraw-Hill, New York (2002)

Concept of Dynamical Traps: Model Systems of Human Actions and Experimental Evidence

Ihor Lubashevsky, Arkady Zgonnikov and Dmitry Parfenov

Abstract Dynamical traps as a new emergence mechanism related to the bounded capacity of human cognition is considered. It assumes that individuals (operators) governing the dynamics of a certain system try to follow an optimal strategy in controlling its motion but fail to do this perfectly because similar strategies are indistinguishable for them. This is described in terms of some neighborhood of the equilibrium point, the region of dynamical traps, wherein each point is regarded as an equilibrium one by the operators. So when a system enters this region and while it is located in it, maybe for a long time, the operator control is suspended. A simple model of oscillator with dynamical traps and the characteristic features of its dynamics are discussed. Experiments on the balancing of a virtual pendulum were conducted to examine the basic features of human control over unstable systems that are expected to be affected by human fuzzy rationality. It is demonstrated that practically only the dimensions of the phase space region wherein a given pendulum trajectory is located depend on the subject age and skill as well as the pendulum parameters determining the difficulty of the balancing. In contrast, the forms of the distribution functions are the same for all the subjects. The data of the virtual experiments are compared to the results of numerical simulation of the oscillator with dynamical traps. The phase trajectories and the phase variable distributions are shown to be similar for the two systems. In addition a chain of oscillators with dynamical traps which mimics cooperative interaction of human operators is considered also. It is, actually, demonstrated that the human fuzzy rationality can cause complex cooperative dynamics in many-element ensembles.

I. Lubashevsky (✉) · A. Zgonnikov · D. Parfenov
University of Aizu, Fukushima, Tsuruga, Ikki-machi, Aizu-Wakamatsu965-8560, Japan
e-mail: i-lubash@u-aizu.ac.jp

A. Zgonnikov
e-mail: arkadiy.zgonnikov@gmail.com

D. Parfenov
e-mail: narytyan@gmail.com

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_15,

Keywords Dynamical traps · Human control · Human fuzzy rationality · Pendulum balancing

1 Introduction

Recent advances in the field of human control have given evidence to the fact that humans do not generally operate systems under their control in a precise way. Maintaining a system exactly at the desired position requires the ability of the human operator to keep perfect awareness and to react immediately even to the smallest deviations. Meanwhile, experimental studies have revealed that the considerable response latency and the effects of noise in the sensorimotor system prevent human operators from implementing the continuous control strategies (see, e.g., [1] and references therein). Instead, the discontinuous (intermittent) control is found to be efficient in the presence of time delays and random perturbations in human-controlled processes [2].

This property of human behavior is a manifestation of the phenomenon often referred to as human fuzzy rationality [3]. Up to now there have been a few attempts to develop a mathematical formalism capturing the effects of human fuzzy rationality. In particular, the basic model of human reaction threshold is commonly used in applied studies, but still there is a lot of uncertainty about the intrinsic mechanisms causing the anomalous behavior of systems under human control (see, e.g., [4]). The dynamical trap model [5–7] being a certain version of the fuzzy threshold concept is another alternative to the standard fixed-point attractor in complex socio-psychological systems. Both of the two models capture the human fuzziness in driving a system towards the desired end-state. They introduce a certain region around the hidden equilibrium, where each state is treated as acceptable by the operator.

In the present work we compare the results of numerical simulation of a simple oscillator with dynamical traps and experimental data collected in balancing a virtual pendulum to find out some general features caused by the human fuzzy rationality. In addition a chain of such oscillators mimicking cooperative behavior of several human operators in controlling a complex system is considered.

Dynamical Trap Model

To illustrate the dynamical trap concept, let us consider the following model of oscillator in the phase space $\{x, v = \dot{x}\}$

$$\dot{v} = -\Omega(v)\omega^2(x + \sigma v) + \epsilon\xi(t)\,, \tag{1}$$

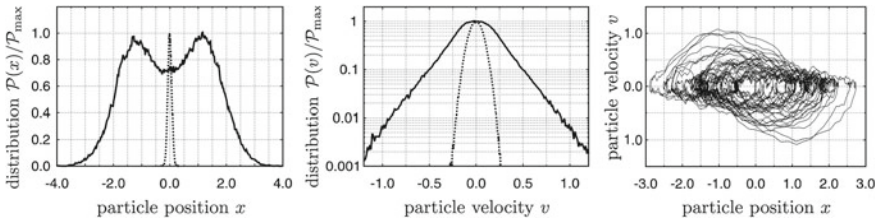


Fig. 1 The distribution functions and the phase portrait for the oscillator with dynamical traps governed by Eq. (1), based on results of [6]. The used parameters are $\sigma = 1$, $\Delta = 0.01$, $\epsilon = 0.1$, and the space-time scales were chosen such that $\omega = 1$ and $v_{\text{th}} = 1$. The *dotted lines* match the absence of dynamical traps, $\Delta = 1$

where x is the position of the particle and v is its velocity. This model imitates the dynamics of a system driven by a human operator to the stationary point $x = 0$, $v = 0$ [7]. Here the ansatz

$$\Omega(v) = \frac{\Delta v_{\text{th}}^2 + v^2}{v_{\text{th}}^2 + v^2} \tag{2}$$

such that $\Omega(0) \ll 1$ and $\Omega(v) \approx 1$ for $|v| \gg v_{\text{th}}$ describes the stagnation of the operator actions in the region $\mathbb{Q}_t = |v| \lesssim v_{\text{th}}$ because in this case he just does not know what to do since all the points in the region $\mathbb{Q}_t \ni \{0, 0\}$ seem equivalent to him. Here v_{th} may be regarded as a fuzzy threshold of the operator perception of the system velocity. The parameter $\Delta \ll 1$ is the measure of dynamical trap strength, the effective friction coefficient $\sigma > 0$ and the frequency ω quantify the operator actions at distant points $\{x, v\}$ when the necessity of driving the system towards the origin becomes well recognizable for him. The white noise $\xi(t)$ with the amplitude ϵ mimics the actions of uncontrollable factors.

The characteristic features of oscillations described by model (1) are demonstrated in Fig. 1. They are: (i) the bimodal distribution $\mathcal{P}(x)$ of the particle position x, (ii) the Laplace form of the velocity distribution $\mathcal{P}(v)$, and (iii) the form of the phase portrait shown in Fig. 1 (right frame). Besides, the system dynamics may be treated as a sequence of alternating fragments of motion outside the region $\mathbb{Q}_t$ and inside it. The action points where, in fact, the operator halts or resumes his control over the system motion are effectively regarded as smooth transitions between the regular and stochastic regimes [7].

2 Virtual Pendulum Balancing

We have analyzed the process of balancing an inverted virtual pendulum embedded into viscous environment. The task of human stick balancing has been investigated widely from various perspectives; studies based on both real-world and virtual experiments are available (see, e.g., [8, 9]). However, up to now, attention was paid mainly

to the in-depth understanding of the mechanical and psycho-motor aspects of human control. The purpose of the present work is to elucidate some general properties of human cognition. It regards the pendulum balancing as just a characteristic example of dynamical processes controlled by humans. The effect of viscosity endowing the pendulums with the over-damped dynamics allows us to appeal to the model of oscillator in the two dimensional space $\{\theta, \dot{\theta}\}$. Here the variable θ is the angle between the vertical axis and the line passing through the center of mass of the corresponding pendulum, the variable $\dot{\theta}$ is its angular velocity. It should be noted that without human actions the dynamics of over-damped oscillators is described by the first order differential equations and the corresponding phase space contains only one variable, the angle θ. However, the operator is able only to move the cart in controlling the pendulum state. As a result the cart velocity becomes an additional phase variable, which extends the system phase space and prompts us to regard the angle velocity $\dot{\theta}$ as an independent phase variable. In other words, the system dynamics under human control is determined not only by the stick angle but the angular velocity as well.

The motion of the analyzed virtual pendulum embedded into viscous environment is described by the following dimensionless model:

$$\tau\dot{\theta} = \sin\theta - \frac{\tau}{l}\, v(t)\cos\theta\,. \tag{3}$$

Here τ is a time scale characterizing the pendulum fall without human control, l is the characteristic pendulum size, and v is the current velocity of the cart motion governed by the mouse movement. Equation (3) stems from the moment balance of all the forces acting on a similar real pendulum embedded into highly viscous liquid. It should be noted once more that the phase space of the given system has to comprise not only angle θ but also its derivative $\dot{\theta}$. This assumption is due to the fact that the operator controlling the system evidently perceives the angular velocity of the stick and regulates the value of control effort $\upsilon(t)$ based on the current values of both factors.

In order to conduct the virtual experiments, we developed a simple software tool that implements the model described above. The operator has to maintain the angle between the virtual stick and the vertical axis near the unstable equilibrium position $\theta_{eq} = 0$ by moving the cart via computer mouse.

Experiments

It has been found that for all the subjects the statistical properties of the pendulum motion controlled by human actions are notably similar. Namely, the portraits formed by irregular trajectories of the pendulum motion on the "angle–angular velocity" plane and the distribution functions of the corresponding phase variables are of high similarity for all the subjects.

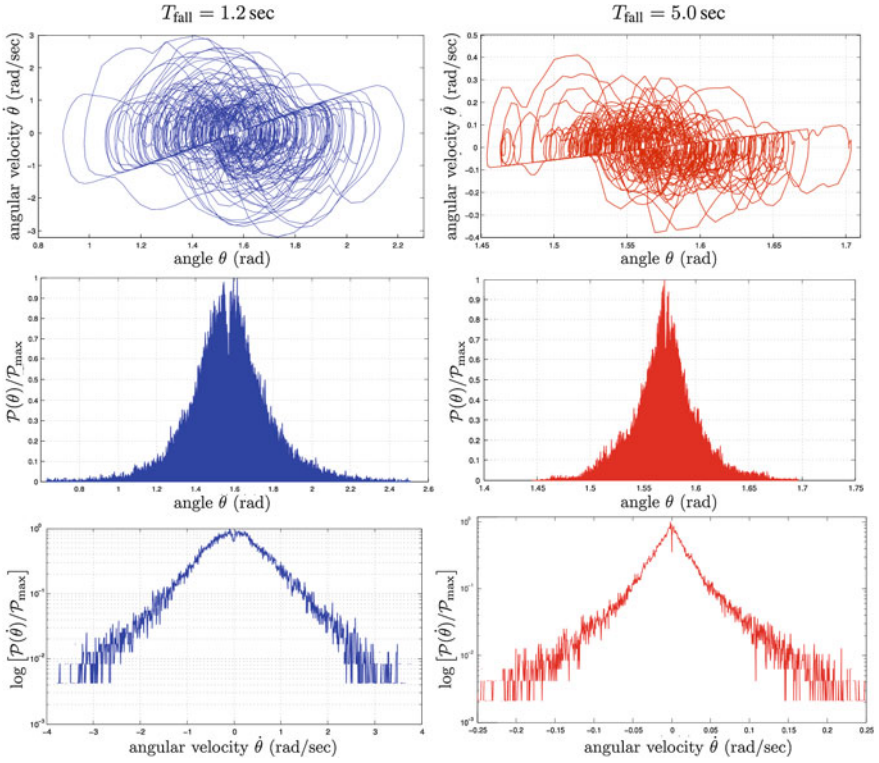


Fig. 2 The distribution functions of the angle θ, the angular velocity $\dot{\theta}$, and the phase portraits obtained during balancing the inverted stick for different values of the fall time T_{fall}. Hear the pendulum angel is measured with respect to the horizontal line

Let us discuss some of the found results in more details. The main parameter of system (3) determining the pendulum behavior is the time scale τ that may be interpreted as the value characterizing the fall time T_{fall} of the pendulum not affected by the operator. Empirically we found the following relation between these values: $T_{\text{fall}} \approx 4\tau$ s. We figured out that for $T_{\text{fall}} < 1$ s all the subjects failed to perform task most of the times. For $T_{\text{fall}} > 4$ s the subjects demonstrated perfect performance, i.e., there were no registered failures in performing the task. In order to figure out the dependency of the system phase portrait and the distribution functions of the phase variables on the time scale parameter, we analyzed the data obtained during the experiments for various values of T_{fall}. The results are represented in Fig. 2. It is clearly seen that for two values of T_{fall} that are close to the opposite boundaries of the feasibility domain the phase portraits and the variable distributions are self-similar in structure. For T_{fall} in between these two values the results remain similar. Similar to the oscillator with dynamical traps (i) the distribution function $\mathcal{P}(\theta)$ of the pendulum angle θ is bimodal, (ii) the distribution function $\mathcal{P}(\dot{\theta})$ of the angular velocity $\dot{\theta}$ is

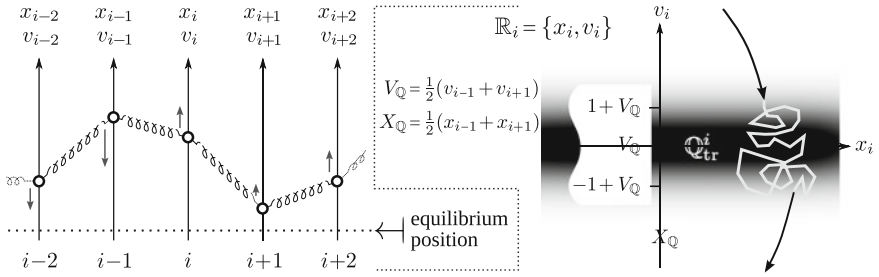


Fig. 3 The chain of N beads under consideration and the structure of their individual phase space $\mathbb{R}_i = \{x_i, v_i\}$ $(i = 1, 2, \ldots, N)$. The formal initial $i = 0$ and terminal $i = N+1$ beads are assumed to be fixed, specifying the equilibrium bead position

of the Laplace form, also (iii) the phase portraits are similar in geometry. Only the characteristic scales are affected substantially by the value of the parameter τ.

3 Lazy Bead Model

As far as the cooperative interaction of human operators is concerned, the following model captures some characteristic features of such human behavior. Let us consider a chain of N "lazy" beads (Fig. 3). Each of these beads can move in the vertical direction and its dynamics is described in terms of the deviation $x_i(t)$ from the equilibrium position and the motion velocity $v_i(t) = dx_i/dt$ depending on time t, here the bead index i runs from 1 to N. The equilibrium position $x_i = 0$ is specified assuming the formal initial $(i = 0)$ and terminal $(i = N + 1)$ beads to be fixed. Each bead i "wishes" to get the "optimal" middle position with respect to its nearest neighbors. So one of the stimuli for it to accelerate or decelerate is the difference

$$\eta_i = x_i - \frac{1}{2}(x_{i-1} + x_{i+1})$$

provided its relative velocity

$$\vartheta_i = v_i - \frac{1}{2}(v_{i-1} + v_{i+1})$$

with respect to the pair of the nearest beads is sufficiently low. Otherwise, especially if bead i is currently located near the optimal position, it has to eliminate the relative velocity ϑ_i, representing the other stimulus for bead i to change its state of motion. The model to be formulated below combines both of these stimuli within one cumulative impetus $\propto (\eta_i + \sigma\vartheta_i)$, where σ is the relative weight of the second stimulus.

When, however, the relative velocity ϑ_i becomes less then a threshold θ, i.e., $|\vartheta_i| \lesssim \theta$, bead i is not able to recognize its motion with respect to the nearest neighbors. Since a bead cannot "predict" the dynamics of its neighbors, it has to regard them as moving uniformly with the current velocities. So from its standpoint, under such conditions the current situation cannot become worse, at least, rather fast. In this case bead i just "allows" itself to do nothing, i.e., not to change the state of motion and to retard the correction of its relative position. This feature is the reason why such beads are called "lazy". Below we will use dimensionless units in which, in particular, the perception threshold is equal to unity $\theta = 1$.

Under these conditions the equation governing the system dynamics is written in the following form

$$\frac{dv_i}{dt} = -\Omega(\vartheta_i)[\eta_i + \sigma\vartheta_i + \sigma_0 v_i] + \epsilon\xi_i(t)\,. \tag{4}$$

If the cofactor $\Omega(\vartheta_i)$ were equal to unity, the given system would be no more then a chain of beads connected by elastic springs characterized by the friction coefficient σ. The term $\sigma_0 v_i$ with the coefficient $\sigma_0 \ll 1$ that can be treated as a certain viscous friction of the beads moving via a medium into which the given system is embedded has been introduced to prevent the beads from attaining extremely high velocities. The factor $\Omega(\vartheta_i)$ is due to the effect of dynamical traps and is given again by ansatz (2). Model (4) allows for random factors in terms of white noise $\xi_i(t)$ affecting the motion of bead i with intensity ϵ.

Results of Simulation

In order to analyze the dynamical trap effect on its own the noise absence case was studied first. The system dynamics was found to depend on the intensity of "dissipation" quantified by the parameter σ. When the parameter σ is not too small the system tends to get the regime of regular dynamics represented by a collection of limit cycles of individual bead motion. It should be noted that these limit cycles could be of complex form when the number of beads is not too large, namely, $N \lesssim 10$ [7]. Nevertheless for systems with large number of beads the resulting phase portrait takes a rather universal form shown in Fig. 4 (left frame). However, the "time to formation" T_N, i.e. the mean time required for a given bead chain to get the steady state regular dynamics grows exponentially as the number of beads increases. For example, for beads with $\sigma = 1$ this time can be approximated by the function

$$T_N \approx T_c \cdot \exp\{N/N_c\} \qquad \text{with} \quad T_c \sim 60 \quad \text{and} \quad N_c \sim 13 \tag{5}$$

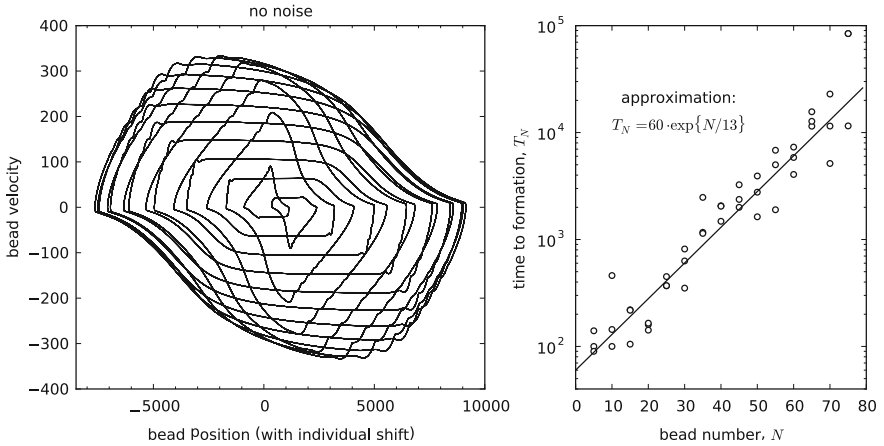


Fig. 4 The characteristic phase portrait of the steady state dynamics exhibited by systems without noise and not too weak "dissipation" (*left frame*). The chain of 30 beads with $\sigma = 1$ was used in constructing the shown pattern where the limit cycles of each second bead are visualized. The *right frame* depicts the characteristic time T_N required for such a system to get the steady state dynamics vs the number N of beads. The *scatted points* are the data obtained for each value of N on three trials, $\sigma = 1$ was used in simulation

(see Fig. 4 (right frame)). It enables us to pose a question about regarding the chaotic dynamics of such systems for $N \to \infty$ as a certain phase state.

Noise forces these systems to undergo two phase transitions as its intensity ϵ increases. The first one can be categorized as the transition from the regular bead motion to a cooperative chaotic bead motion. The latter means that the beads correlate substantially with one another in motion but individual trajectories are rather irregular and the magnitude of this irregularity cannot be due to the present noise only. The second transition is determined by the formation of highly irregular mutually independent oscillations in the bead position. To illustrate the first phase transition Fig. 5 depicts two phase portraits of the middle bead motion for different values of ϵ. As seen, for $\epsilon = 0.01$ the phase portrait looks like a regular limit cycle disturbed by small noise. In contrast, when the noise intensity increases by two times, i.e., $\epsilon = 0.02$, the corresponding phase portrait becomes rather complex in form and the volume of the phase space layer containing the shown trajectory as a whole sharply grows. Exactly the two features has enabled us to classify the found effect as a phase transitions. It should be noted, that this phase transition from regular motion to stochastic chaos, in contrast to the second transition to highly irregular motion, does not manifest itself in the one-particle distributions of all the variables x, v, η, ϑ ascribed to the beads individually, so, it could be categorized as a "weak" phase transition.

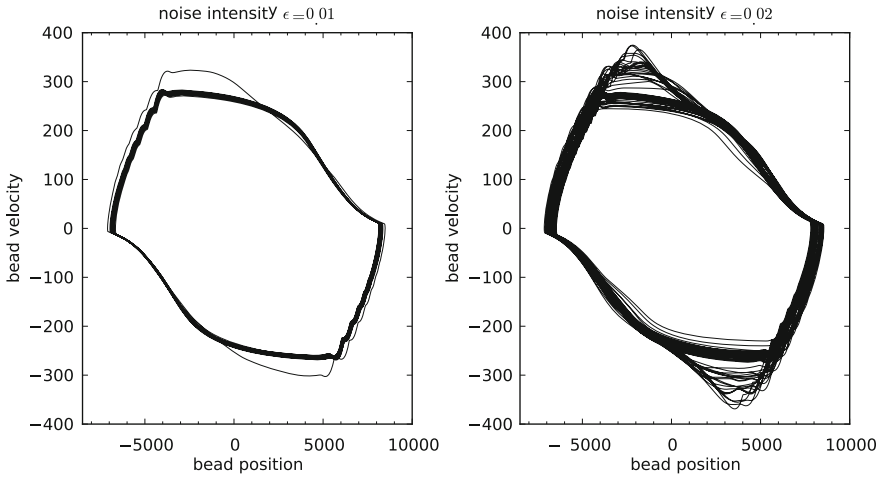


Fig. 5 The phase portraits of the middle bead motion of the 30-bead chain with $\sigma = 1$ for two values of the noise intensity $\epsilon = 0.01$ and 0.02. In plotting these portraits bead trajectories of motion during time interval about 2×10^4 were used

4 Conclusion

Nowadays the fact that the human control of unstable systems is characterized by intermittency, time delay, and the effects of human prediction is well established. However, the basic properties of human control are still rather far from being understood well. The present research has been aimed at finding some universal properties of human control by conducting experiments on balancing virtual pendulums.

The notion of dynamical traps was introduced to describe possible effects caused by the bounded capacity of human cognition in ordering events or actions according to their preference. Its particular implementation is that human beings as active elements of a certain system cannot individually control all the governing parameters within the accuracy required for stabilizing the system dynamics perfectly. Therefore one chooses a few crucial parameters and mainly focuses attention on them. When the equilibrium with respect to these crucial parameters is attained the human activity slows down, retarding in turn the system dynamics as a whole.

We have dealt with a virtual pendulum whose dynamics is assumed to be overdamped, which mimics the balancing in viscous liquid. We have found out that the balancing difficulty, the operator age and skill affect mainly the amplitude of the object motion only. The three features of the operator behavior remain the same; they are (i) the general shape of both the angle and angular velocity distributions, (ii) the two-peak structure of the angle probability density function, and (iii) the structure of the phase portraits formed by the object motion in the phase space “angle - angular velocity”.

The obtained experimental data also contribute to the evidence for the intermittency of human control. The most important, though, is the universality of the distribution functions for all the subjects and all the considered models. This fact could be possibly explained by the effects of learning how to balance a stick with minimal efforts. This fact may enable us to pose a question about the existence of the general laws describing human cognition near the reaction threshold, which must be of probabilistic nature.

Moreover, one may even speculate that the obtained results give evidence to the fact that the standard notion of fixed-point attractor may not be applicable in dynamical systems where human role is crucial due to the phenomena of fuzzy rationality.

By way of example, we considered emergent phenomena in chains of coupled oscillators with dynamical traps. This system was studied numerically. As demonstrated, without noise the system dynamics tends to the regime of regular bead motion if the friction coefficient is not too small. However, the characteristic time required for a given system to get this regime grows exponentially with the number N of beads. It enables us to pose a question about regarding the chaotic transient processes as a certain phase state in the limit $N \to \infty$. When the friction coefficient becomes sufficiently small the steady state dynamics of such systems can undergo transition to chaotic bead motion even for chains with small number of beads. Depending on its intensity noise can induce the formation of three characteristic phases, highly irregular individual oscillations of the beads, the cooperative chaotic bead motion, and the synchronized regular bead motion. It should be noted that the transition between the regimes of regular and cooperative chaotic bead motion manifests itself only the sharp growth of the volume of the phase space layer containing the bead trajectories, whereas all the one-particle distribution functions does not change their forms remarkably.

Acknowledgments The work was supported in part by the JSPS "Grants-in-Aid for Scientific Research" Program, Grant 24540410-0001, and the Competitive Research Funding of the University of Aizu, Project P-21 (FY2013).

References

1. Loram, I.D., Gollee, H., Lakie, M., Gawthrop, P.J.: Human control of an inverted pendulum: is continuous control necessary? Is intermittent control effective? Is intermittent control physiological? J. Physiol. **589**, 307–324 (2011)
2. Cabrera, J.L., Milton, J.G.: On-off intermittency in a human balancing task. Phys. Rev. Lett. **89**, 158702 (2002)
3. Dompere, K.K.: Fuzzy Rationality. Springer, Berlin (2009)
4. Cabrera, J.L., Milton, J.G.: Stick balancing, falls and dragon-kings. Eur. Phys. J. **205**, 231–241 (2012)
5. Lubashevsky, I., Hajimahmoodzadeh, M., Katsnelson, A., Wagner, P.: Noise-induced phase transition in an oscillatory system with dynamical traps. Eur. Phys. J. B **36**, 115–118 (2003)

6. Lubashevsky, I., Mahnke, R., Hajimahmoodzadeh, M., Katsnelson, A.: Long-lived states of oscillator chains with dynamical traps. Eur. Phys. J. B **44**, 63–70 (2005)
7. Lubashevsky, I.: Dynamical traps caused by fuzzy rationality as a new emergence mechanism. Advs. Complex Syst. **15**, 1250045 (2012)
8. Milton, J.G., Ohira, T., Cabrera, J.L., Fraiser, R.M., Gyorffy, J.B., Ruiz, F.K., Strauss, M.A., Balch, E.C., Marin, P.J., Alexander, J.L.: Balancing with vibration: a prelude for "Drift and Act" balance control. PLoS One **4**, e7427 (2009)
9. Suzuki, S., Harashima, F., Furuta, K.: Human Control Law and Brain Activity of Voluntary Motion by Utilizing a Balancing Task with an Inverted Pendulum. Advs. Hum.-Comput. Interact. (2010) Article ID: 215825 (2010)

Model of Cognitive Functions for Description of the Creative Design Process with Computer Support: Improving of the Interpretation Method for the Computer Conceptual Re-Design

Jakub Jura and Jiří Bíla

Abstract This chapter describes model of the human cognitive functions, especially these ones which are important for the creative process. The broad context of this work is a development of the conceptual redesign method with computer support (called CRDP—Computer Redesign Process). This method is based on postmodern principles of the interpretation, on the respect to complexity of the creative process and at the impossibility of its direct control. Psychological approaches (e.g. the mind mapping [1] or creativity timing) are used in this interpretation method and its core is a creating of the interpretation map. The aim of this submission is to describe the emergent design processes for the purpose of their simulation and method's HCI improving. The model is made in Unify Modelling Language. The fractal approach to the communication between user (designer) and software system (CRDP) was outlined.

Keywords Unified modeling language (UML) · Cognitive science · Cognition · Interpretation · Conceptual design · Redesign · Human-Computer interaction

1 Introduction

Creative conceptual design is always nontrivial mental process—likewise its computer support. Design can be conceived of as a purposeful, constrained, decision making, exploration and learning activity [2]. This approach oriented us to the field of Cognitive Science where the human factor play crucial role. The person—designer—in his creative movements is very complex system. Authors designed the method of

J. Jura (✉) · J. Bíla
Department of Instrumentation and Control Engineering, CTU, Prague, Czech Republic
e-mail: Jakub.Jura@fs.cvut.cz

J. Bíla
e-mail: Bila@vc.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_16,

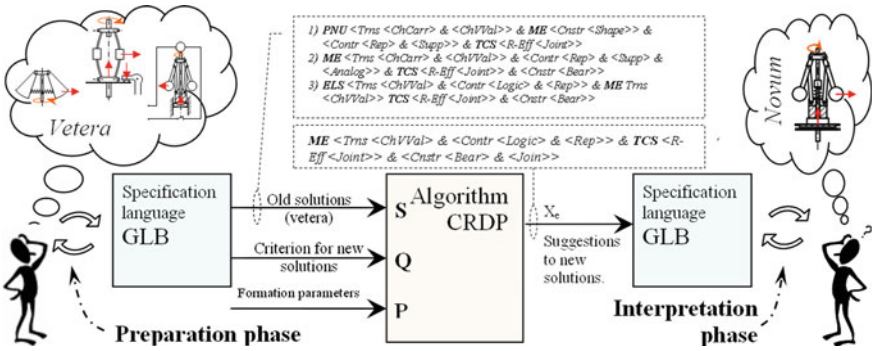


Fig. 1 Description of designing process with CRDP software and methodical support

computer support of the creative conceptual redesign process which is based on the principles of **emergence in the interpretation process** [3, 4]. And here suggested UML cognitive model facilitates a development of the mentioned method.

1.1 Computer and Methodical Support of Conceptual Design

There are many algorithms, methods and procedures (like a TRIZ/ARIZ or Morphological analysis) for the facilitating of the synthesis of the innovative concepts. One of them is **Computer ReDesign Process** (CRDP), which was developed on Faculty of Mechanical Engineering of the CTU in Prague (e.g. [5]). Inputs to the CRDP software system (algorithm CRDP on the Fig. 1) are (1) three old solutions (**vetera**), (2) criterions for a new solution and (3) formation parameters (fields of activities and principles which form a new solution).

The output is a set of suggestions to an innovation (**novum**). The old and new solutions are described in a specification language GLB [6, 7]. Global Context (**GLB**) is a language, which conceptualizes the domain of the conceptual design and represents semantic properties of knowledge elements by means of pre-formed semantic structures. System CRDP is equipped by the interpretation of their outputs. This interpretation method helps user's creative processes to find out the solution [3, 4] (Fig. 1).

2 Knowledge Based UML Model

First step of our work was to create the part of the model, which represents selected psychological pieces of knowledge. And they are fundament for continue modeling of the emergent creative processes.

2.1 UML

Unified Modeling Language [8] is standard primarily developed for software engineering, but it is possible to use them for the description of any systems in general. UML is a graphical language and modeled subject is possible to describe by thirteen types of diagram. We utilize only three of them: Class diagram for the description of the structure properties, State diagram for the description of the class's behavior and Sequence diagram for the description of interactions between classes. Unified Modeling Language was already used for modeling of the cognitive functions at the field of **transportation engineering** by [9]. They used UML for modeling of the driver attention and described it by the state diagram. Our approach to **cognitive modeling** is not usual at the field of **engineering psychology** (and psychology in general).

2.2 Lexical Analysis and Model

The **skeleton** of the UML model arises from the process of the **lexical analysis** (e.g. [10]). Problem field of human cognition was described in nature language (e.g. Sternberg and Sternberg 2012) and names of the classes were derived from the nouns (and nouns phrases) used in the text after the selection (the selection rules are part of the OMT lexical analysis). In the similar way were obtained the names of the attributes from the adjectives. And associations and operations were obtained from the verbs and verbal phrases.

2.3 Structure of the Knowledge: Class Diagram

The statical structure of the described system is represented by the UML class diagram. This view comes out from the lexical analysis and on this basis was made the skeleton of this model (V1).[1] However this one contains uncovered logical spaces (**inconsistency**) which has been resolved by the **addition of the connecting pieces of knowledge**. This improved model is possible to see at Fig. 2. (And is mark as V2).

It is possible to assume that the model will be improved by its next use (e.g. Chap. 3). Model can be enriched with the model of reality (in this case veterum, novum, system CRDP, GLB language etc.) and its mental representation. There is opening questions about borders of the modeled system.

[1] V1 … the first version.

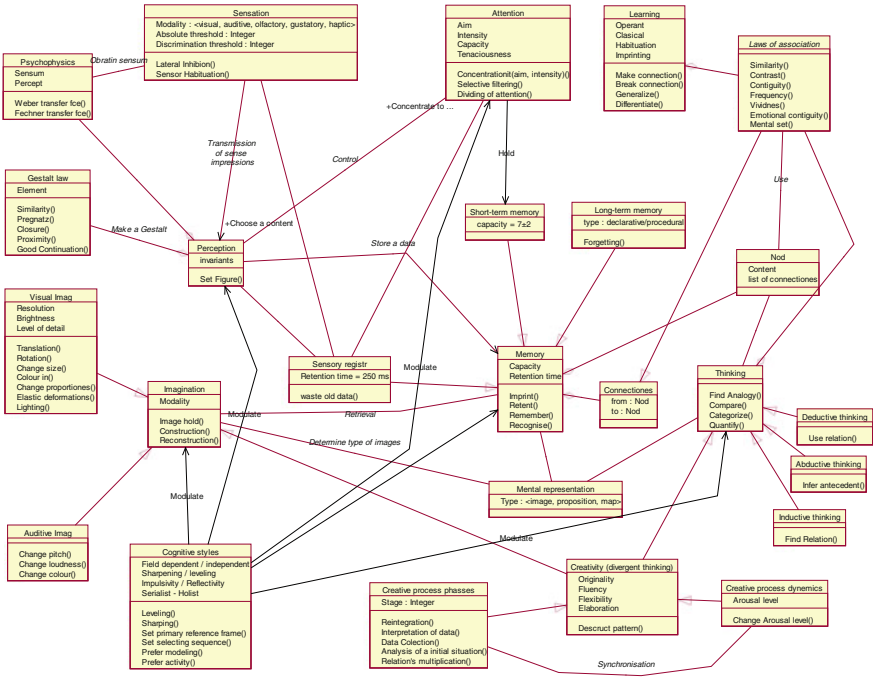


Fig. 2 Structure of the cognitive functions described by the UML Class diagram V2

2.4 Behavior of the Class "Creative Process": State Diagram

The state diagram is used for the description of the behavior of the given class. State diagrams were made for few key classes only. One of them is the class "Creative process phases" (see Fig. 3). **Emergence of a new solution** is possible to model as a violence of the structure invariant of the common transformation (e.g. [11, 12]).

Creativity phases are here related to the arousal[2] of the organism and genomic expression. We suppose, that the **arousal cycle** (Fig. 4—bold line), which has a **fractal character**, is possible to derive from the communication interaction (which is also fractal—Fig. 6). For the description of the communication process is used UML sequence diagram (Fig. 5). Description of these phases was derived from the Deep neurobiology of E. Rossi [13].

[2] Our conception of the "arousal" is different from Rossi and is comprehended as a unipolar psychophysiologic variable related to the Ascendant Reticular Activation System (ARAS).

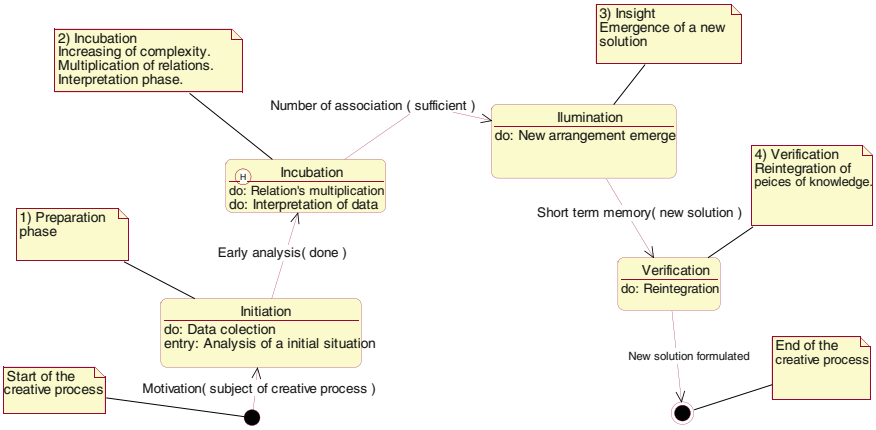


Fig. 3 State diagram of the class Creative process phases

3 Description of the CRDP Interpretation Method by the UML Sequence Diagram

Sequence diagram notice a **communication between classes**. Since the UML model of cognitive functions is oriented to the redesign process, so the sequence diagram view the **process of the communication between user (designer) and CRDP software** primarily (V1).

3.1 Incremental Improving Loop: Impact to the Model

First version of the model V0 was built on three classes only—Designer, CDRP interface and interpretation map. This model was improved and the internal designer structure was deeper elaborated. Also veterum and novum description was added (see Fig. 5).

Considered next improve of the model suppose an elaboration of the psychological processes and interconnections between the state diagram of the creative process class and the mentioned sequence diagram. Especially the fractal communication patterns should be considered. Unfortunately the UML standard does not support it.

4 Fractal Property of Real Communication Act

Ideal dyadic communication act is simple exchange (communication action and reaction … as it is consider e.g. in UML sequence diagram). The basic unit of it we can call the **big exchange**. But real communication act is fractal process. Big exchange precedes smaller exchange and it precedes smaller etc. (as it is shown on Fig. 6). It is easy to observe this phenomenon at the example of human communication

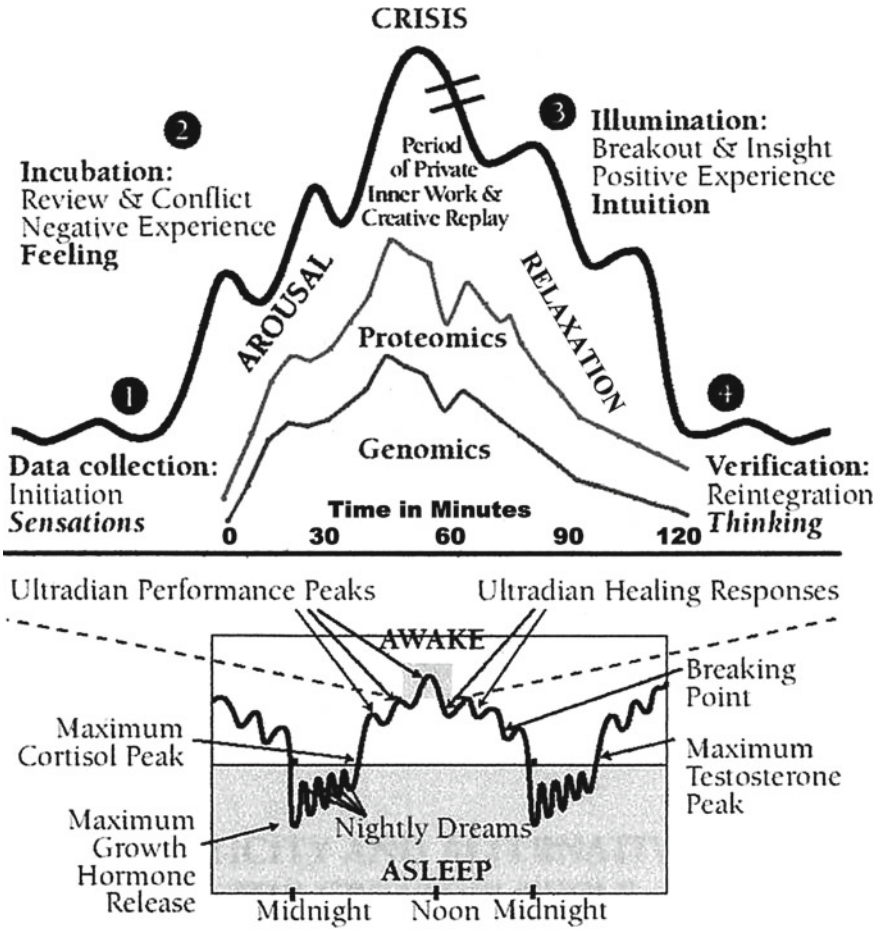


Fig. 4 Description of the development of the creative process related to the arousal of the organism (*Source* [13])

or in a nature [14]. Preceding small act prepare the communication context at the both sides.

We have the hypothesis that the high level of the synchronization of the internal or external communication "**punctuation**" (term by [15]) leads to the **emergence of novelty**. This synchronization can cause effect like a resonation and can significantly affect the human arousal process, which [13] related to the creative process (Chap. 2.4).

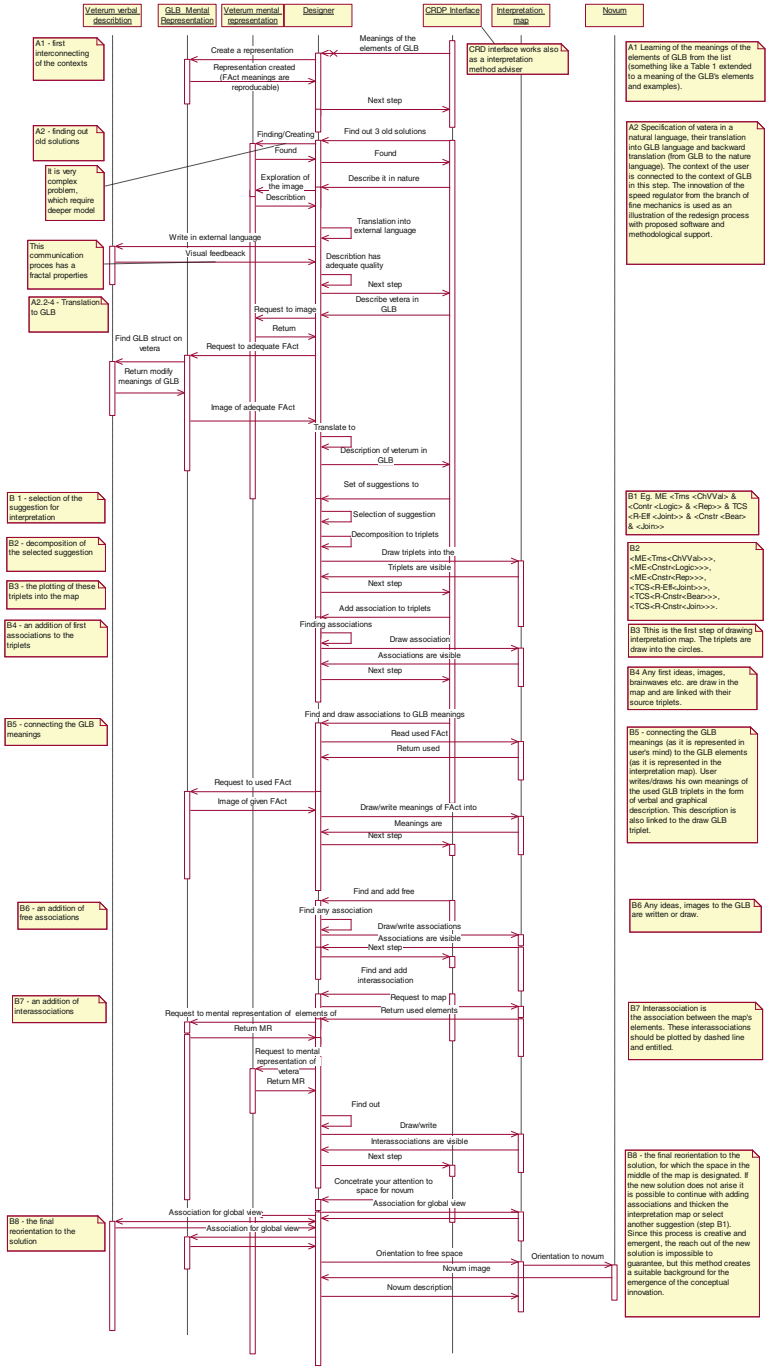


Fig. 5 Sequence diagram of the communication between user (designer) and CRDP software (V1)

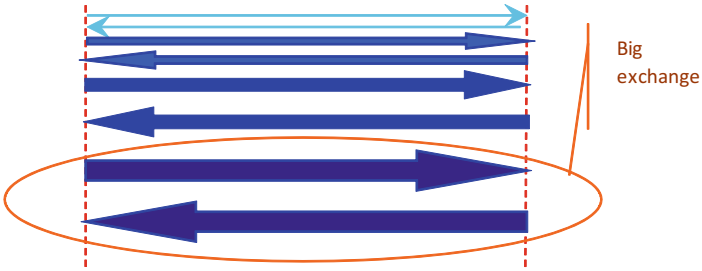


Fig. 6 Schema of the fractal communication pattern

5 Meta-Concepts Background

Usually we construct theories in the way of the reduction of the complexity. Meanwhile in the case of the description of the creative process (or divergent thinking) this assumption is not useful. If we want to help user in his creative task, we have to fully respect his nature internal processes and this processes are very **complex**, **irreducible** and has **fractal character**.

5.1 2nd Order Cybernetics

One of the theoretical approaches, which respects principles above is 2nd order cybernetics—cybernetics of cybernetics [16]. We used principle of the circular causality and suppose the autopoiesis [17] at the level of emergence of a new creative solution.

5.2 Postmodern Approach

Following approach have very strong connections to the 2nd order cybernetics, but is more philosophically oriented. Our approach respects the plurality of points of view, impossibility of direct control or finally the whole process in which new solution arises is analogy to Maturana's "languaging" [18]. The object modeling principles supports this approach.

6 Conclusion

The first and second version of the psychological model of the designer cognitive process was created. The model is expressed in the graphical language UML. One of objectives of this work is to develop mentioned design method CRDP—especially their human-computer interface. At the base of the model is possible to fit proposed CRDP interpretation method to their user (their psychological limits and modes

of work). Moreover, this submission is a step to integration (or reintegration) of psychological knowledge from the field of cognition and creativity (in a given context). The submission also outlines the course for the further work on this issue.

Acknowledgments The development of this chapter has been supported by Research Grant SGS12/177/OHK2/3T/12. This support is very gratefully acknowledged.

References

1. Buzan, T.: The Ultimate Book of Mind Maps. Harper Collin Publisher, London (2005)
2. Gero, J.S.: Creativity, emergence and evolution in design: concepts and framework. Knowl.-Based Syst. **7**(9), 435–448 (1996)
3. Jura, J., Bila, J.: Interpretation Method for Software Support of the Conceptual Redesign Process: Emergence of new concepts in the interpretation process. In: International Conference on Education and research in Computer Aided Architectural Design in Europe, pp. 227–233, FA CVUT, Prague, Czech. Rep. (2012)
4. Jura, J.: Interpretation process in conceptual re-design of systems, PhD thesis, Faculty of Mechanical Engineering of the CTU in Prague, Prague, Czech. Rep. (2012)
5. Bila, J., Tlapak, M.: Knowledge Discoveries and Emergent Synthesis in Conceptual ReDesign Process. International Conference on Computational Intelligence for Modelling. Control and Automation—CIMCA'05, pp. 537–543. Austria, Vienna (2006)
6. Bila, J., Jura, J., Tlapak, M.: Emergent Synthesis in Conceptual ReDesign Process. In: Proceedings of 6th International Workshop on Emergent Synthesis—IWES' 06, Kashiwa, Japan (2006)
7. Bila, J., Tlapak, M.: Ontologies and Formation Spaces for Conceptual ReDesign of Systems. In: 10th International Conference on Advanced Engineering Design—AED′04, E2, 1–8 CD ROM, Glasgow, Scotland (2004)
8. OMG—Object Management Group Inc: Unified Modeling Language: UML Resource Page, http://www.uml.org (c1997)
9. Haring, K.S., Ragni, M., Konieczny, L.: A Cognitive Model of Drivers Attention. In: 11th International Conference on Cognitive Modeling, Germany, Berlin (2012)
10. Schalley, C.A.: Cognitive Modeling And Verbal Semantics: a Representational Framework Based On UML. De Gruyter Mouton, Berlin (2004)
11. Bila, J.: Algebras of transformations in the detection of unexpected situations of UX3 type. In: 16th Interantional Conference on Soft Computing. pp. 495–500, University of Technology, Brno, Czech. Rep. (2010)
12. Bila, J.: Detection of emergent situations by structural invariants. In: 17th International Conference on Soft Computing, pp. 534–539, Brno, Czech. Rep (2011)
13. Rossi, E.L., Rossi, K.L.: The New Neuroscience Of Psychotherapy, Therapeutic Hypnosis and Rehabilitation: a Creative Dialogue With Our Genes. In: Milton, H. (ed.) Erickson Institute of the California Central Coast, Los Osos CA (2008)
14. Jura, J., Bila, J.: Computation of the Fractal Dimension of Meteorological Quantities. In: Proceedings of 16th International Conference on Soft Computing—Mendel 2010, pp. 140–145, Brno University of Technology, Brno, Czech. Rep. (2010)
15. Watzlawick, P., Beavin-Bavelas, J., Jackson, D.: Pragmatics of Human Communication—A Study of Interactional Patterns. Pathologies and Paradoxes. W. W, Norton, New York (1967)
16. von Foerster, H.: Cybernetics of Cybernetics. University of Illinois, Urbana Illinois (1974)
17. Maturana, H.R., Varela, F.: Autopoiesis and Cognition: the Realization of the Living. Reidel Publishing Company, Dordecht (1980)
18. Maturana, H.R.: Biology of Cognition. Biological Computer Laboratory, Research Report BCL 9.0. Urbana IL, University of Illinois. (1970)

Dynamical Systems Approach to Atherosclerosis Modeling

Johan L. A. Dubbeldam

Abstract Mathematical modeling of clinical systems is difficult as these usually comprise complex systems with many interacting components. We show how it is still possible to model these systems by making use of a dynamical systems point of view. By calculating bifurcation diagrams, one can discriminate between different models and clinical parameter regimes can be identified. The emphasis in this presentation will be in particular on models of atherosclerosis, but the suggested approach is applicabe to a much wider class of clinical models.

Keywords Dynamical systems · Modeling · Bifurcation analysis · Atherosclerosis

1 Introduction

Making mathematical models for the development of certain diseases is a tedious effort in which contacts between people performing experiments and mathematicians are indispensible. However, even then it turns out that many parameters in the system are not known or vary greatly from person to person. It would therefore be a great advantage to be able to discriminate between models without knowing the detailed parameters. Furthermore, it would be beneficial to have a clue from the mathematical model in what range physical parameters are supposed to take values.

We illustrate here how to obtain results from mathematics that can be helpful in a clinical environment by giving some examples of predictions of a model that may be verified in practice. Finally, we discuss how mathematical models consisting of ordinary differential equations discussed here could be extended to systems of partial differential equations in which the interaction between the many components are much more involved.

J. L. A. Dubbeldam (✉)
Delft University of Technology, Delft, The Netherlands
e-mail: j.l.a.dubbeldam@tudelft.nl

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_17,

2 Model

The model (evolution equations) are given as:

$$\dot{m} = \left(\frac{aL}{(1+\sigma)(1+L)} - \epsilon - c \right) m, \tag{1a}$$

$$\dot{M} = cm - \frac{bML}{1+L}, \tag{1b}$$

$$\dot{L} = \frac{dm}{f+m} - eLM - L, \tag{1c}$$

$$\dot{F} = \frac{bLM}{1+L}. \tag{1d}$$

The physical interpretation of this coupled system is described in [1] and we will only give a brief explanation of the terms and parameters here. All parameters and variables in the model are dimensionless and nonnegative. The dot denotes the time derivative and since the equation for F (1d) decouples from the system, we only consider the three equations for m, M and L. In this model, which is called model A, in accordance with Ref. [1], the shear stress σ is simply considered a parameter. The systems describes how the components that constitute a plaque that resides in the artery evolve in time. The meaning of the parameters and variables is as follows. The monocytes are denoted by (m). They evolve due to conversion into macrophages (M) at a rate c, diffusion out of the plaque region and they are recruited from the blood due to signalling of oxidized LDL molecules (L). The macrophages arise from conversion of monocytes, but are in turn converted to so-called foam cells by the ingestion of LDL. The LDL concentration is determined by the ingestion of macrophages and the recruitment due to oxidation of LDL molecules that migrate from the blood to the plaque region where they are oxidized almost immediately by the monocytes. In this model the foam cells, form a residue which remains after the macrophages have ingested LDL and die. This model which captures a number of important aspects of the progression of atherosclerosis for long times will next illustrate how clinical consequences can be derived from even a qualitative model like this.

We first remark that when studying the equilibria of Eqs. (1a–1d) it can be easily seen that the system contains two classes of equilibria. The first class is of the following form $\{(0, M_0, 0) | M_0 > 0\}$ and contains infinitely many equilibria. The second class of equilibria consists of two points $(m^{\pm}, M^{\pm}, L^*)$. The expressions of $m^{\pm}$, $M^{\pm}$, and $L^{\star}$ can easily be calculated or found in [1] and will not be reproduced here.

We should further remark that the system (1d) does not exhibit choatic behavior for the range of parameters that we investigated. It turns out that the saddle point has a two-dimensional attracting manifold that extends to infinity and therefore there the systems is approximately two-dimensional and chaos can be ruled out. For chaotic phenomena in blood(vessels) we refer to [2, 3]

2.1 Bifurcation Analysis

In order to obtain results about the dynamics of the system Eqs. (1a–1d) we perform a bifurcation analysis. It turns out to be advantageous to not simple perform this analysis in one parameter, but rather in two parameters. In this way we obtain that the bifurcation diagram is organized by a so-called Bogdanov-Takens point, which denotes a point where the system has two eigenvalues zero. From bifurcation theory [4] is then known that two curves, one curve of Hopf bifurcations and a curve of limit cycle bifurcations emerge from this point. We can now illustrate the dynamics in the system by studying the different regions of the bifurcation diagram. Of course, an important question is which parameters should be chosen for the bifurcation diagram. The intake of cholesterol modeled by parameter d is an obvious control parameter. The second choice is less straightforward. We found that the choice of b leads to some interesting results. First of all it allows curves in the bifurcation diagram to be calculated analytically. Second, as b models the life time of the macrophages it gives insight in the time scale of the biochemical processes in atherosclerosis. In fact we find that atherosclerosis can be prevented by a very low intake of cholesterol and, more surprisingly, a sufficiently long life time of the macrophages.

This can be understood as follows. If the macrophages live longer then they will ingest more LDL particles. In this way the macrophages impede the growth of LDL and consequently the plaque growth will settle down to a stable value. A shorter life time of the macrophages on the other hand leads to larger production of LDL, which in its turn make the monocyte concentration grow, leading eventually to an increased macrophage concentration. Because of the intricate interplay of the components a full bifurcation analysis is needed to predict the dynamical behavior. The full bifurcation diagram is published elsewhere [5], but the dynamics of the different regions in the bifurcation diagram are displayed in Fig. 1.

Because the full dynamics of the system is now known it is possible to compare our results with clinical measurements to see if one of the parameter regions fits the data and so the model can be validated. Using such a validation we can progress by developing a partial differential equations model.

3 PDE Model for Atherosclerosis

The growth of plaque in an artery is, of course, much more complicate [6, 7] than can be described by a simple system of ODEs. One should typically take into account that σ in Eqs. (1a–1d) is not a parameter, but depends on the flow and the shape of the plaque. We started to construct a model for such a system in which the plaque is actually treated as a moving boundary. This problem is complicated as it involves many different time scales and the blood flow contains in general some recirculation regions. We will not reproduce the Navier-Stokes equations here and the coupling to the plaque equations. We will illustrate the consequences of the flow profile on the

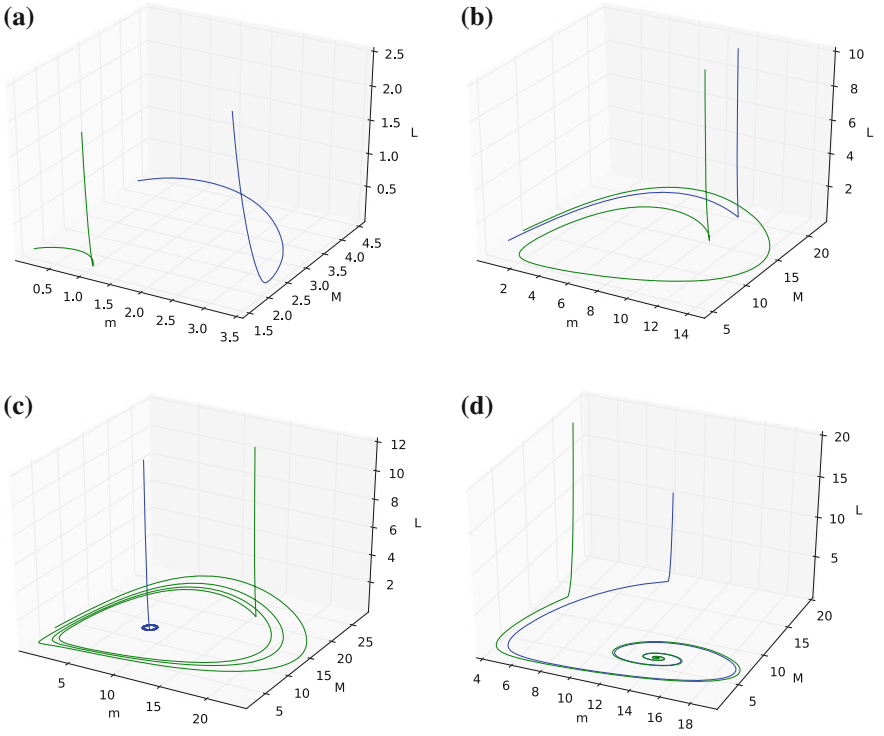


Fig. 1 The four different types of dynamics that arise from varying the parameters b and d. In the first phase diagram **a** there are only type I equilibria, and therefore the M-axis is a globally attracting manifold. In **b** the system has gone through a saddle node bifurcation, but the type II equilibria are unstable. In **c** the system exhibits bistability. There is an unstable limit cycle and a stable equilibrium of type II. In **d** the unstable limit cycle has disappeared and depending on the initial conditions, the system arrives at the invariant line $(0, M_0, 0)$ or in (m^+, M^+, L^*)

plaque growth by Fig. 2 in which the plaque and the flow profile is shown at different moments in time. Of course, the model that we implement should still give the same results as the ODE model for the limiting case of a plaque that is smeared out over the artery wall. For different plaque shapes, however, which are usually also found in experiments, the plaque growth can vary considerably.

4 Conclusion

We have illustrated using a simple model for atherosclerosis how bifurcation techniques can help constructing models for clinical practice. Using simple ODE models, validation is possible, after which a more elaborate (PDE) model can be constructed.

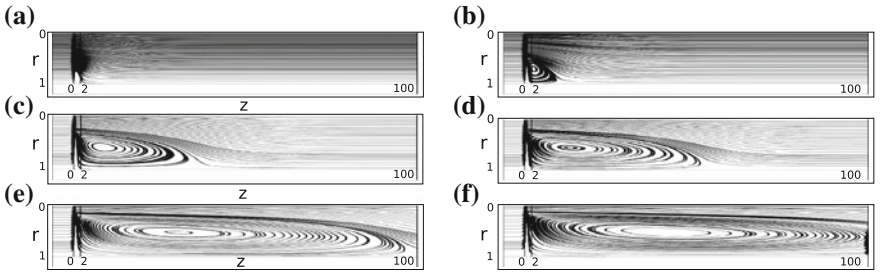


Fig. 2 The streamlines vary enormously in time due to the response of the growth of the plaque in the artery. In **a** the stenosis has just started and in **f** an enormous recirculation zone is present. To take this into account in the model, numerical simulation with time dependent grids are necessary or one has to resort to asymptotic methods

Moreover, we have illustrated how bifurcation theory can help to identify parameter ranges which are consistent with experiments.

References

1. Bulelzai, M.A.K., Dubbeldam, J.L.A.: Long time evolution of atherosclerotic plaques. J. Theor. Biol. **297**, 1–10 (2012)
2. Schelin, A.B., et al.: Chaotic advection in blood flow. Phys. Rev. E. **80**, 016213 (2009)
3. Schelin A.B., et al.: Fractal structures in stenoses and aneurysms in blood vessels. Phil. Trans. Roy. Soc. A. **368** (2010)
4. Kuznetsov, Y.A.: Numerical normalization techniques for all codim 2 bifurcations of equilibria in ODE's. SIAM J. Numer. Anal. **36**, 1104–1124 (1999)
5. Bulelzai, M.A.K., Dubbeldam, J.L.A., Meyer, H.: Submitted to Physica D. (2013)
6. Gijsen, F.J.H., Wentzel, J.J., Thury, A., Mastik, F., Schaar, J.A., Schuurbiers, J.C.H., Slager, C.J., van der Giesen, W.J., de Feyter, P.J., van der Steen, A.F.W., Serruys, P.W.: Strain distribution in human coronary arteries relates to shear stress. Am. J. Physiol. Heart Circ. Physiol. **295**, H1608–H1614 (2008)
7. Ross, R.: Atherosclerosis: a defense mechanism gone awry. Am. J. Pathol. **143**, 987–1002 (1993)

Deterministic Modeling Spatio-Temporal Dynamics of Delay-Induced Circadian Oscillations in *Neurospora crassa*

Dmitry Bratsun and Andrey Zakharov

Abstract We propose a spatially extended deterministic model with time delay for the circadian oscillations in the fungal species *Neurospora crassa*. The temporal behavior of the system is governed by a two variable model based on the nonlinear interplay between the *FRQ* and *WCC* proteins which are products of transcription of *frequency* and *white collar* genes. We show numerically that the model accounts for various features observed in experiments. Spatio-temporal protein patterns excited in *Neurospora* in complete darkness are studied for different initial conditions. It is shown that basal activation of transcription factors has a strong effect on pattern formation.

Keywords Time-delay · Circadian rhythms · Neurospora · Pattern formation

1 Introduction

Circadian rhythms are biological rhythms that are common to almost all living organisms. A remarkable feature of these rhythms is that they are not simply a response to 24 h environmental cycles imposed by the Earth's rotation, but instead are generated internally by cell autonomous biological clocks. After the decades of research, the genetic mechanism of circadian oscillations has been widely recognized as a core of this phenomenon. As it is known now, a feedback influence of protein on its own expression can be delayed which leads to non-Markovian phenomena in this system [1]. It is evident that the delay prevents the system from achieving equilibrium, and results instead in the familiar limit cycle oscillations. The deterministic

D. Bratsun (✉) · A. Zakharov
Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia
e-mail: dmitribratsun@rambler.ru

A. Zakharov
e-mail: az1211@mail.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_18,

Fig. 1 Conidiation of *Neurospora* growing on solid agar medium

and stochastic properties of gene regulation taking into account the non-Markovian character of gene transcription/translation were studied in [2, 3].

In fact, a filamentous fungus *Neurospora crassa* (hereinafter *N.c.*) is an excellent model system for investigating the mechanism of circadian rhythmicity because of the wealth of genetic and biochemical techniques available. *N.c.*'s rhythms of asexual spore-formation produces easily assayed bands of conidia spores in cultures growing on solid agar medium (Fig. 1). A number of mutations are available that affect circadian rhythmicity, and molecular analyses of some of these genes have contributed to models for the circadian oscillators that are currently thought to be applicable to many other organisms. No wonder that since the rhythm was discovered in the *N.c.* [4], this organism has become a research polygon in the declared area.

With advances in molecular biology, understanding of *N.c.*'s circadian clock has improved, and main genetic components of this clock have been determined [5].

In this work, we propose the model of circadian oscillations in the *N.c.* which is a further simplification of models proposed by Smolen et al. [6] and Sriram and Gopinathan [7]. In this paper we have used the deterministic description of the system focusing on its spatial dynamics. In the part of spatially extended delay-induced circadian oscillations our modeling seems to be first in the literature. Perhaps this can be explained by a prevailing tradition in the study of circadian oscillations. It is possible also that this is due to computational difficulties that arise when studying reaction-diffusion systems with time delay. To overcome these difficulties, we have proposed a new method for the numerical study of such systems [8].

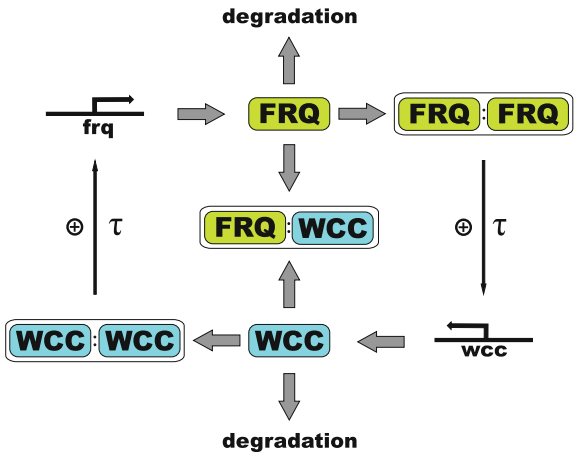


Fig. 2 Network architecture of the circadian rhythm molecular components in *Neurospora*

2 Two Variable Model of Circadian Rhythms in Neurospora

The simplified graphical depiction of protein network architecture responsible for circadian rhythms is presented in Fig. 2. The primary molecular components of the circadian oscillator are the *frequency* (*frq*) and *white collar* genes (*wc-1* and *wc-2*) which form a feedback loop comprised of both positive and negative elements [5]. The white-collar proteins *WC-1* and *WC-2* are transcription factors which form a heterodimeric complex known as the *WCC* complex. The *WCC* acts as a positive regulator of *FRQ* by activating its transcription in the dark, while the frequency protein dimerizes and then acts as a negative regulatory element by binding to and inhibiting the function of *WCC*. As the circadian cycle progresses the *FRQ* protein is phosphorylated and degraded which allows the cycle to begin anew. Also, these species can be removed by association with *WCC* to form *FRQ/WCC* complexes. Furthermore the production of *WC-1* and *FRQ* proteins are subject to a delay on the order of several hours. The previous experimental efforts have highlighted also the importance of degradation of the core clock components particularly that of *FRQ*, plays in establishing the period of the circadian rhythm.

In our model we assume there are two primary components: the heterodimeric *WCC* complex and the *FRQ* protein. We start our analysis from a set of biochemical reactions constituting the mechanism of bioclock

$$F + F \xrightarrow{k_1^F} F_2, \quad W + W \xrightarrow{k_1^W} W_2, \tag{1}$$

$$F_2 \xrightarrow{k_{-1}^F} F + F, \quad W_2 \xrightarrow{k_{-1}^W} W + W, \tag{2}$$

$$D_0^F + W_2 \xrightarrow{k_2^F} D_1^F, \quad D_0^W + F_2 \xrightarrow{k_2^W} D_1^W, \tag{3}$$

$$D_1^F \xrightarrow{k_{-2}^F} D_0^F + W_2, \quad D_1^W \xrightarrow{k_{-2}^W} D_0^W + F_2, \tag{4}$$

$$D_1^F(t) \xrightarrow{k_F} D_1^F + F^{t+\tau_F}, \quad D_1^W(t) \xrightarrow{k_W} D_1^W + W^{t+\tau_W}, \tag{5}$$

$$F \xrightarrow{B_F} \varnothing, \quad W \xrightarrow{B_W} \varnothing, \tag{6}$$

$$F + W \xrightarrow{k} \varnothing, \tag{7}$$

$$\varnothing \xrightarrow{A_F} F, \quad \varnothing \xrightarrow{A_W} W, \tag{8}$$

where F and W stand for number of isolated monomers of *FRQ* and *WCC* respectively. In (1)–(8) the reactions of dimerization and dedimerization are given by (1) and (2), the transitions between operator-site states for each protein due to binding and unbinding some dimer to the promoter of corresponding gene are reflected by (3) and (4), the time-delayed productions of corresponding proteins is given by (5), the linear and non-linear degradation are described by (6) and (7), the basal transcriptions are given by (8) respectively. By supposing that reactions of dimerization (1)–(2) and binding/unbinding (3)–(4) are fast in comparison with the processes of production/degradation of proteins (5)–(8), we can conclude that their dynamics quickly enters into a local equilibrium and arrive finally to the following two-variable spatially extended system (the procedure for derivation of the equations is not presented here, see [9] for more details):

$$\frac{\partial F}{\partial t} = \frac{1}{1+4K_1^F F}\left(A_F + k_F \frac{K_1^W K_2^F W^2(t-\tau)}{1+K_1^W K_2^F W^2(t-\tau)} - B_F F - kFW\right) + D\left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}\right), \tag{9}$$

$$\frac{\partial W}{\partial t} = \frac{1}{1+4K_1^W W}\left(A_W + k_W \frac{K_1^F K_2^W F^2(t-\tau)}{1+K_1^F K_2^W F^2(t-\tau)} - B_W W - kFW\right) + D\left(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2}\right). \tag{10}$$

Here D is the coefficient of protein diffusion in the cell. For simplicity, we assume that the diffusion coefficients of *FRQ* and *WCC* proteins are equal. Even supposing that the delay is defined by the length of the path traveled by RNA polymerase along the gene, one obtains different values since the *wc-1* gene (it is part of the locus NCU02356.5) is a one and a half times longer than the *frq* gene (it is in the locus NCU02265). But the exact values of the delays are currently unknown, and for simplicity we assume that time delays are equal to the same value τ. As one can see, the model (9)–(10) includes a positive feedback loop in which activation of *FRQ* production by *WCC* increases the level of *FRQ*, leading to an increase in the level of WCC itself. The negative feedback loop in which *FRQ* represses the *frq* gene transcription by binding to the *WCC* is also modeled. Generally, the temporal part of model (9)–(10) is a further simplification of models proposed by Smolen et al. [6]

and Sriram and Gopinathan [7]. Despite of its simplicity, the equations (9)–(10) still reflect all the characteristic features of circadian rhythms in *N.c.*

In (9)–(10) we also have taken into account a basal transcription governed by rates A (8). In eukaryotes, an important class of transcription factors called basal or general transcription factors (GTFs) are necessary for transcription to occur. Many of these GTFs do not actually bind DNA but are a part of the large transcription preinitiation complex that interacts with RNA polymerase directly.

As it is known, the *N.c.* not only has the advantage that powerful genetics and molecular techniques are able to be performed on it, but it has another advantage—the circadian rhythms of conidiation are easily monitored on Petri dishes. To observe the phenotypic expression of the clock, conidia are inoculated at some place of a Petri dish. After growth for a day in a constant light, the position of the growth front is marked and the culture is transferred to constant dark. The light-dark transfer synchronizes the cells in the culture and sets the clock running from subjective dusk. Following transfer, the growth front is marked every 24 hours with the aid of the red light, which has no effect on the clock. The growth rate is constant and the positions of the readily visualized orange conidial bands (separated by undifferentiated mycelia) allow determination of both period and phase of the rhythm. Thus, the computational domain $\Sigma \in (x, y)$ where the protein fields are solved numerically can be interpreted as a flat area of two-dimensional physical space of a Petri dish occupied by the mycelium of *N.c.* In fact, *N.c.* is a multicellular organism, and the translation of proteins occurs within individual cells. But we can consider the mycelium of the fungus as a whole due to the important feature of *N.c.*: the mycelium of the organism consists of branched hyphae which show apical polar growth. The fungal hyphae are typically composed of multiple cells or compartments demarcated by septa with the central pore sometimes up to $0.5\,\mu\mathrm{m}$ in diameter. Thus, the protein produced in the separate cells of *N.c.* seems to be able to cross the intercellular walls, and we can assume an existence of joint molecular cloud of protein inside a whole organism.

In order to perform two-dimensional simulations of circadian oscillations governed by Eqs. (9) and (10), we define the domain $\Sigma{:}(0 < x < 200, 0 < y < 200)$ with zero-flux boundary conditions for concentrations of *FRQ* and *WCC* proteins. The initial-boundary value problem (9)–(10) has been solved by a finite difference method was described in detail in [8]. The explicit scheme was adopted to discretize equations. The equations have been approximated on a rectangular uniform mesh 400×400 using a second order approximation for the spatial coordinates.

3 Spatio-Temporal Nonlinear Dynamics

In the numerical calculations we have used the following fixed values of parameters: $K_1^F = 5$, $k_F = 8\,\mathrm{nM/h}$, $K_1^W = 5$, $k_W = 4\,\mathrm{nM/h}$, $K_2^F = 5$, $B_F = 0.3\,\mathrm{h}^{-1}$, $K_2^W = 5$, $B_W = 0.4\,\mathrm{h}^{-1}$, $k = 30\,\mathrm{nM}^{-1}\,\mathrm{h}^{-1}$, $\tau = 6\,\mathrm{h}$, $D = 0.01\,\mathrm{m}^2\,\mathrm{s}^{-1}$. These values have been taken from [6], plus we add some our own, e.g. for dimerization and diffusion.

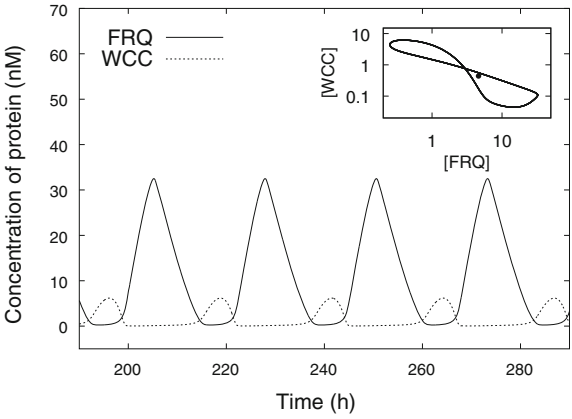


Fig. 3 The time series and phase portrait (*inset*) of *FRQ* (*solid line*) and *WCC* (*dashed line*)

In Fig. 3, we plot the time series and phase portrait of the total number of *FRQ* and *WCC* proteins obtained by solving nonlinear equations (9)–(10) for zero basal transcription. The period of oscillations is about 22.65 h. Above the Hopf bifurcation point, in the phase space there is an unstable stationary point bounded by a stable limit cycle (Fig. 3, *inset*). All trajectories are attracted to the periodic solution. No subcritical oscillations have been found, and the transition seems to occur smoothly.

Let us discuss now the results of numerical simulation of spatially extended model (9)–(10). Two different cases of initial conditions will be considered. In the first case shown in Fig. 4, the initial state of the system was the random distribution of the *FRQ* and *WCC* proteins over the entire area of integration. It corresponds to a situation where each node initially has its own phase of oscillations. At time $t = 0$, all these rhythms are independent, but in the process of evolution they enter into the nonlinear interaction, which results in different types of spatio-temporal behavior. From the point of view of biology, a random distribution of initial phases is somewhat artificial, since a synchronization of biorhythms in cells occurs at the stage of embryonic development of the organism. Nevertheless, the numerical study of the system with random initial conditions allows developing a deeper understanding of the potential of the system to self-organize and to evaluate its probable forms of pattern formation.

Since *FRQ* and *WCC* are always in anti-phase, we can choose only one of them to illustrate the system dynamics. Figure 4 presents the patterns formed by the concentration of *FRQ* protein shown for four consecutive points of time. As it can be seen from figure, the nonlinear dynamics of spatially extended system consists of two distinct oscillatory modes. One is the oscillations occurring in spatially ordered interacting cells (Fig. 4, $t = 500$). The characteristic size of the cells increases slowly as time goes on. Thus, this is a typical quasi-standing wave pattern. The second oscillatory mode is a spiral traveling-wave pattern (Fig. 4, $t = 3000$). These waves arise from selected initial disturbances (Fig. 4, $t = 500$). Each spiral wave travels outward in all directions from its source. It continues until the spiral wave pattern occupies the entire domain Σ. Note that if the front of wave looks more or less orderly, by

entering deeply inside the secondary instability area there have appeared numerous secondary centers of the excitation of the spiral waves. The nonlinear interaction between them leads to the formation of chaotic pattern (Fig. 4, $t = 4100$).

Thus, the evolution of the system passes through two stages: first, a slowly time-varying cellular structure has appeared, which synchronizes the oscillations of different nodes. Since the system is far away from equilibrium, the fluctuations give rise to several spiral traveling waves which immediately break up in the core inducing the spatiotemporal chaos. Even though the level of the protein oscillates randomly in space, the system is in standby mode. We will show below that some external or internal stimuli can synchronize the system in space and time.

Let us consider the second case of initial conditions, when evolution of the system starts from a small local perturbation of *FRQ* (or *WCC*) against zero field in the remaining part of the domain Σ. It was established experimentally that the front of mycelium propagates radially in all directions with the constant velocity starting from the location where it was placed initially. It should be emphasized that the local phase of the oscillations caused by the regular concentric traveling wave is determined by the phase of the initial perturbation in the bud, and the spatio-temporal pattern as a whole is formed during the morphogenesis of the organism.

Thus, the oscillations in the bud can be referred to as a "global clock" of Neurospora. Figure 5 presents the pattern after 125 hours of the evolution (left) and the one-dimensional cross-section of the protein field (right). We see that *N.c.*'s conidia have formed six bands that correspond to the period of oscillations about 22.65 h (compare with spatial pattern shown in Fig. 1).

4 Effect of Basal Transcription on Pattern Formation

The interesting properties of the circadian rhythms reflected in our model, are the synchronization and disappearance of oscillations with an increase of the basal production rate of *FRQ* or *WCC*, which occur, for example, under constant light conditions. Figure 6 shows the various dynamic regimes in the parameter plane spanned by two rates of basal transcription. The Hopf bifurcation curve divides the parameter plane into the part where the stable periodic solution exists (inside the balloon), and the part where there is stable steady state (outside the balloon). One can see that when the rate exceeds some critical value, the amplitude of oscillations becomes zero (Fig. 6, *inset*).

It should be noted that our model predicts that the period of oscillations remains robust to increases in A while the amplitude of oscillations evidently does not. It is interesting to note that this behavior directly contradicts the predictions made by some conceptual models directly correlating period length to the amplitude of oscillations in the core molecular clock components. This fact could be crucial for the experimental proof which of mechanisms is really responsible for oscillations in the *N.c.*

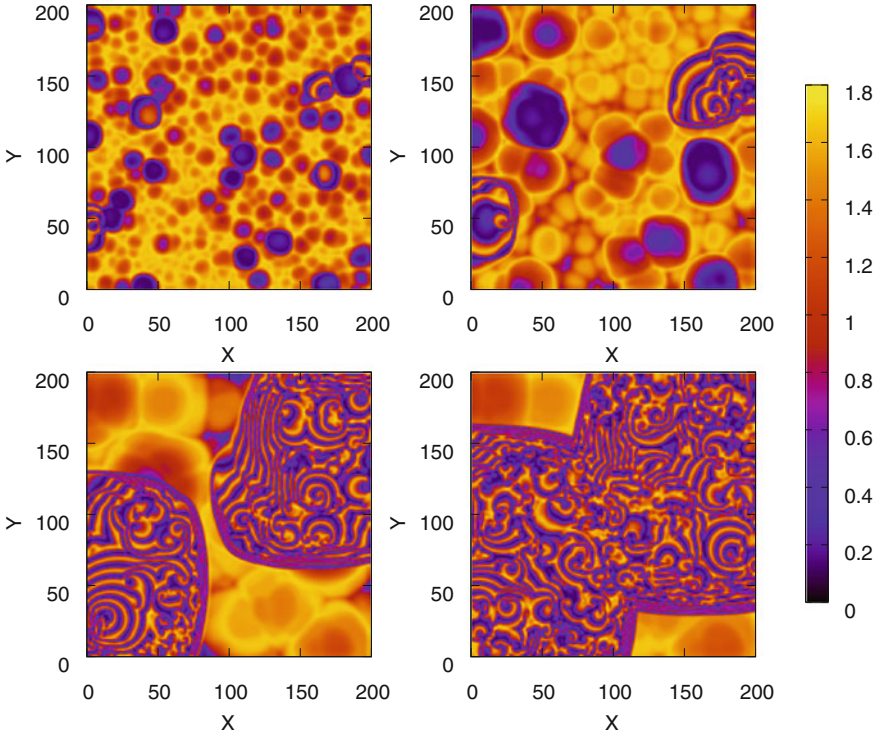


Fig. 4 The evolution of *FRQ* concentration without the basal transcription. The frames from *left* to *right* and from *up* to *down* correspond to times $t = 500, 1200, 3000, 4100$ respectively

The spatial phase synchronization is the process when spatially distributed cyclic signals tend to oscillate with a repeating sequence of relative phase angles. We have noticed above that some external stimuli can synchronize the spatio-temporal behavior of the system. The external control of this active medium can be performed, for example, via the basal transcription factors. We found that system dynamics is particularly sensitive to the basal transcription of the *WCC* protein. Figure 7 presents the density plots of *FRQ* concentration at time $t = 5000$ for four different values of A_W and $A_F = 0$. With increase of *WCC* produced via the basal transcription machinery, the spatio-temporal structure of the system becomes more ordered. In contrast to the distinct chaotic pattern at $A_F = A_W = 0$ (Fig. 7a) formed due to break up of spiral waves, the structures shown in Fig. 7b–d for $A_W = 1; 2; 2.8$ respectively are frozen in space, i.e. each point oscillates periodically in accordance with its phase imposed by spatial pattern. The period of oscillations increases slightly with growth of the rate of basal transcription A_W. We found that further increase of A_W causes a sharp rise in *FRQ* synthesis and the termination of oscillations. It corresponds to a crossing of the Hopf bifurcation curve indicated in Fig. 6.

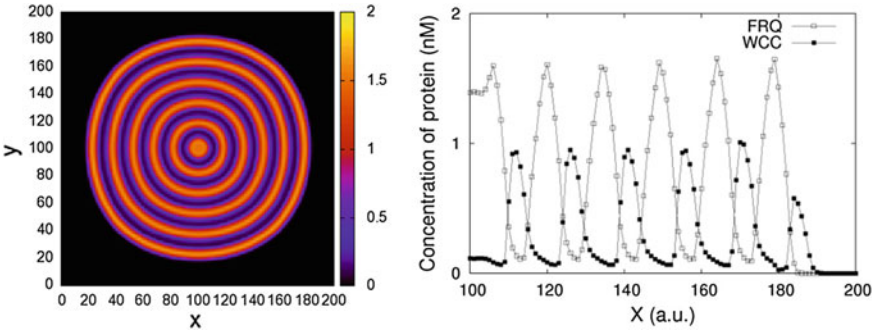


Fig. 5 The density plot of *FRQ* protein after 125 h of evolution developed from the initial disturbance in the center (*left*); transverse profiles of *FRQ* (*open squares*) and *WCC* (*black squares*) proteins at time $t = 125$ (*right*)

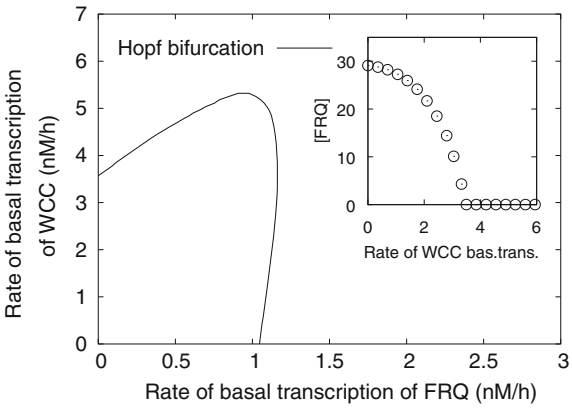


Fig. 6 The neutral curve for the Hopf bifurcation (*solid line*) shown in the parameter plane spanned by two rates of basal transcription. The bifurcation diagram showing the amplitude of oscillations of *FRQ* versus A_W for $A_F = 0$ is plotted in the *inset*

5 Conclusions

The spatially extended deterministic model for delay-induced circadian oscillations in the fungal species Neurospora crassa has been proposed in the paper. The model is suitable to analyze both temporal and spatial dynamics of this organism. The core of the model describing the temporal behavior of the system is based on the interplay between two dynamical variables, concentrations of the *FRQ* and *WCC* proteins. Time delay in the processes of transcription/translation is, perhaps, the easiest source of oscillations in the genetic systems. It exists, apparently, because of the slow movement of RNA polymerase along the gene. It is important to note also that the gene processes are not just slow but also are compound multistage reactions involving the sequential assembly of long molecules. Thus, these processes should obey Gaussian statistics with a certain characteristic mean delay time.

We have shown that our model of circadian oscillations based on the time delay mechanism in transcriptional regulation accounts for various features observed in

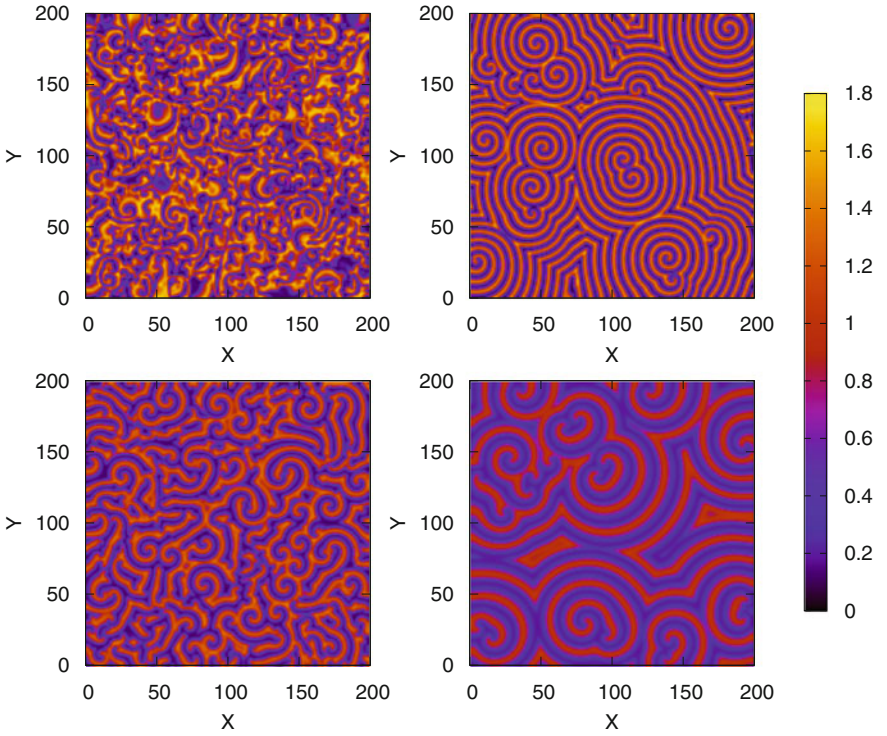


Fig. 7 Density plots of the concentration of the *FRQ* protein at time $t = 5000$ for $A_F = 0$ and different values of A_W: **a** 0; **b** 1; **c** 2; **d** 2.8

experiments such as effect of basal transcription factors, spatio-temporal synchronization and robustness to parameter variation.

Acknowledgments The work was supported by the Department of Science and Education of Perm region (project C26/244), the Ministry of Science and Education of Russia (project 1.3103.2011) and Perm State Pedagogical University (project 031-F).

References

1. Liu, Y., Loros, J.J., Dunlap, J.C.: Phosphorylation of the Neurospora clock protein FREQUENCY determines its degradation rate and strongly influences the period length of the circadian clock. Proc. Natl. Acad. Sci. USA **97**, 234–239 (2000)
2. Bratsun, D., Volfson, D., Hasty, J., Tsimring, L.S.: Delay-induced stochastic oscillations in gene regulation. Proc. Natl. Acad. Sci. USA **102**, 14593–14598 (2005)
3. Bratsun, D.A., Volfson, D.N., Hasty, J., Tsimring, L.S.: Non-Markovian processes in gene regulation. Proc. SPIE **5845**, 210–219 (2005)

4. Pittendrigh, C.S., Bruce, V.G., Rosenzweig, N.S., Rubin, M.L.: A biological clock in Neurospora. Nature **184**, 169–170 (1959)
5. Lakin-Thomas, P.L., Brody, S.: Circadian rhythms in microorganisms: new complexities. Annu. Rev. Microbiol. **58**, 489–519 (2004)
6. Smolen, P., Baxter, D.A., Byrne, J.H.: Modeling circadian oscillations with interlocking positive and negative feedback loops. J. Neurosci. **21**, 6644–6656 (2001)
7. Sriram, K., Gopinathan, M.S.: A two variable delay model for the circadian rhythm of *Neurospora crassa*. J. Theor. Biol. **231**, 23–38 (2004)
8. Bratsun, D., Zakharov, A.: Adaptive numerical simulations of reaction-diffusion systems with history and time-delayed feedback. In Sanayei, A., Zelinka, I., Rössler, O.E. (eds.) ISCS 2013. Emergence, Complexity and ComputationPrague, vol. 8, pp. 1–11. Springer, Heidelberg (2014)
9. Bratsun, D., Zakharov, A.: Modeling spatio-temporal dynamics of circadian rhythms in *Neurospora crassa*. Comput. Res. Model. **3**, 191–213 (2011) (Russian)

Adaptive Numerical Simulations of Reaction-Diffusion Systems with Time-Delayed Feedback

Dmitry Bratsun and Andrey Zakharov

Abstract A new algorithm for calculating the dynamics of spatially-extended reaction-diffusion systems where the current state depends on the whole or partial previous evolution of the system is proposed. The algorithm is based on a finite difference method and involves an adaptive optimization of data storage by storing in a computer memory not all previous nodal data, but only some selected of them, called the base states. The intermediate states are restored by interpolation between the base states. The use of this technique allows the numerical calculations to be implemented on computer systems without large RAM memory. The algorithm efficiency is shown in three numerical examples.

Keywords Time-delay · Finite difference method · Adaptive algorithm for data storage · Reaction-diffusion systems

1 Introduction

Dynamical systems with delay are abundant in nature. They occur in a wide variety of physical, chemical, engineering, economic and biological systems and their networks [1]. The mathematical description of delay dynamical systems will naturally involve the delay parameter in some specified way. This can be in the form of differential equations with delay or difference equations with delay or even might include integral forms. A differential equation with delay describing a dynamical system belongs to the class of retarded functional differential equations [2]. In the present chapter we are focusing on a numerical solution of spatially extended dynamical systems with a

D. Bratsun (✉) · A. Zakharov
Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia
e-mail: dmitribratsun@rambler.ru

A. Zakharov
e-mail: az1211@mail.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_19,

constant discrete delay or history term. Such problems arise in many areas of science ranging from fluid mechanics [3] to complex systems of biological nature [4].

Among the methods of solving delay partial differential equations the method of lines is the most common [5]. These numerical techniques are composed of two consecutive steps. In the first step, the partial derivatives with respect to the spatial direction are replaced by some approximations. At this stage one can apply the finite difference methods [6], the Galerkin finite elements method [7] or some pseudo-spectral method [8]. In the second step, the resulting semi-discrete systems are integrated in time. These systems are composed of stiff ordinary delay differential equations.

It should be noted that the method of lines works well only for one-dimensional equations. If one needs to simulate the equations in more than one spatial direction, one has to use more sophisticated numerical methods (see, for example, [9]). Moreover, if there are non-local terms in the system, this method also cannot be applied.

Generally speaking, the most common numerical method for spatially extended systems is a finite difference method. The explicit finite difference scheme allows it to be implemented as a time recursion which computes a solution at current time from solution samples at earlier times. The simplest explicit scheme requires knowledge of data fields at only one previous time step. However, in order to find the solution at current time to spatially extended system with delay it is necessary to store the data of all nodal data within the range of delay time. In practice, the implementation of this approach makes sense only for low dimensional systems with small delays. Otherwise, the amount of operated data, and hence the calculation time, become so large that it is impossible to carry out a complete numerical study of the system.

In this paper we propose the adaptive finite-difference method for evolutionary reaction-diffusion problems which involves an optimization of data storage.

2 Numerical Method for Evolutionary Reaction-Diffusion Problems with Time Delay

Let us consider the system of nonlinear kinetic equations of diffusion type:

$$\frac{\partial \mathbf{U}}{\partial t} = f(\mathbf{U}(t), \mathbf{U}(t-\tau)) + D\Delta\mathbf{U}, \tag{1}$$

where $\mathbf{U}(x, y, z)$ is the set of physical (or chemical) variables that define the nonlinear kinetics in the absence of diffusion, Δ is the Laplacian operator in Cartesian coordinates x, y and z. The nonlinear function f in (1) consists of both conventional and time-delay terms. For simplicity, we assume that the time delay τ is fixed, although it is not principal for what follows. The system (1) must be supplemented by initial and boundary conditions for $\mathbf{U}$. Several dynamical systems in biology,

optics, economics, ecology, etc., can be described by delay differential equations of the form (1).

The numerical study of spatially extended dynamical system is usually performed using the finite difference method. This method is conventionally considered to be the simplest numerical method for solving partial differential equations. The explicit scheme has a distinct advantage, since it is the easiest to implement and the least numerically intensive coupled with its reliability. Unfortunately, it has one significant drawback: there is a strict limitation on the time step to ensure the numerical stability. Nevertheless, the popularity of the explicit scheme to researchers only grows with time, as the modern analysis of nonlinear dynamical systems is increasingly becoming not just a search for the stationary solution of the system, but also the study of its nonlinear dynamics. In addition, the high-performance computing systems can compensate for the disadvantages of the explicit method. Yet, two of the key issues that arises in computer simulations, are delays in transmitting data between processor to memory and small processor cache memory. These architecture problems manifest themselves clearly in the numerical solution of delay differential equations.

Let the spatial domain Σ be discretized by three orthogonal regular grids based on cubes with spatial subdivisions of h:

$$x_i = ih, \quad i = \overline{0, N_x}; \quad y_j = jh, \quad j = \overline{0, N_y}; \quad z_k = kh, \quad k = \overline{0, N_z}, \qquad (2)$$

whereas time is subdivided into intervals of θ. Each node of the grid (2) represents a certain region, and its value of $\mathbf{U}$ is a measure of the average value of the region. The aggregate of nodal points at fixed time we refer as a time slice. If the system (1) is characterized in terms of a nodal network, one needs to work with an approximate form of the equations. We rewrite the system (1) in the finite difference form using central differences for spatial derivatives and forward difference for time derivative:

$$\mathbf{U}^{n+1}_{i,j,k} = \mathbf{U}^{n}_{i,j,k} + \theta(f(\mathbf{U}^{n}_{i,j,k}, \mathbf{U}^{n-L}_{i,j,k}) +$$

$$+\frac{D}{h^2}(\mathbf{U}^{n}_{i+1,j,k}+\mathbf{U}^{n}_{i-1,j,k}+\mathbf{U}^{n}_{i,j+1,k}+\mathbf{U}^{n}_{i,j-1,k}+\mathbf{U}^{n}_{i,j,k+1}+\mathbf{U}^{n}_{i,j,k-1}-6\mathbf{U}^{n}_{i,j,k})), \quad (3)$$

where L is the number of time steps performed within the range of time delay τ. As it can be seen from (3), to calculate the field at time $n+1$ one needs to know the data not only in the previous moment of time n, but also at time $n - L$ in the past. If the time lag τ is large, the number L is also large. For example, for the numerical integration of delay differential equations in [10] authors have used the forward-Euler method with storage of all delayed quantities for later calculations. They have noticed that integration time steps were reduced until no significant difference was seen after further reduction. The final step size was about 10^{-5}. By taking into account large value of delays (5–10), one can conclude that they were forced to store the data for about a million time steps! It is obvious that such a straightforward approach can not be applied to spatially extended systems with delay since it would demand unrealistically large computer resources.

The explicit method is known to be numerically stable and convergent whenever

$$\theta < \frac{h^2}{2 \max |\mathbf{U}|}. \tag{4}$$

If the integer number of time steps does not fit into the time delay, then the field at current time cannot be constructed. In this case, one must either generate the missing field by interpolating the two nearest neighbors, or fix the time step to a predefined value. In the latter case, the delay should be a multiple of time step (as in [9]).

To resolve all these difficulties, we propose to store the nodal data of not all time slices, but only some of them, which we call as base slices (Fig. 1). The values of variables for the intermediate time slices are restored then by interpolating the data of the base slices, stored in the computer memory. Depending on the smoothness of **U** and the grid dimension, the various types of interpolation can be used. The second idea of the algorithm is a usage of the optimal number of base time slices. It is assumed that their allocation density within the range of time delay τ may vary (Fig. 1). This density supposed to be determined by the smoothness of the function **U** that allows controlling the accuracy of calculations. This approach also reduces error when a variable-step solver dynamically adjusts the time step size.

Let us define the function

$$G(t+\theta) = \max\left(\frac{|U(t+\theta,x,y,z) - U(t,x,y,z)|}{U(t,x,y,z)}\right), \\ (x = \overline{0, N_x}), \quad (y = \overline{0, N_y}), \quad (z = \overline{0, N_z}), \tag{5}$$

which characterizes the relative rate of change of **U**. Then we introduce the important parameter K, which means that each K-th time slice in the data stream within delay interval should be considered as a base slice. When the function **U** changes its value rapidly, one should use all the data within interval of delay: $K = 1$. In fact, it corresponds to the storage of the entire data stream. For reaction-diffusion problems considered in the present work, we have determined empirically that this case occurs at $G > 0.1$. For smaller values of $G(t)$ the algorithm can be used without significant loss of accuracy. In order to minimize the interpolation error and to provide a sufficiently high rate of calculations, the parameter K should take value from the interval [1, 20].

The greatest positive effect of the method is found to manifest itself if the function (5) changes very slowly ($G < 0.01$). In this case, we can store just only every 20th time slice. For larger values of K one cannot meet the specified accuracy of the calculations. Thus, one can use the empirical rule

$$K = \frac{\tau}{\theta(1 + \exp^{(2-50G)})}, \tag{6}$$

for finding the optimal number of base slices. It allows to specify the allocation density of base time slices depending on time-varying function $G(t)$ (Fig. 1).

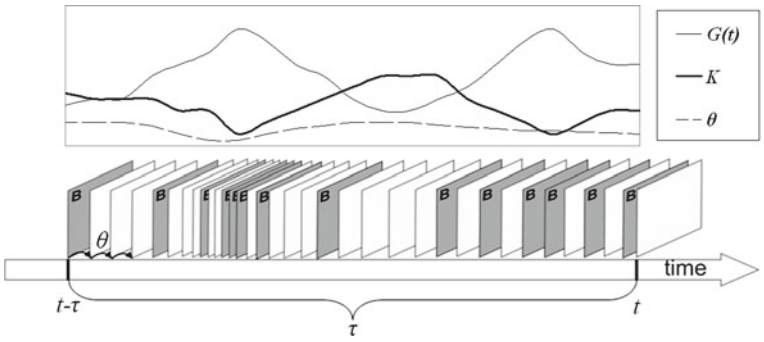


Fig. 1 Schematic diagram of the adaptive method of data storage. Each time slice represents a nodal network at the corresponding time step. The base time slices are indicated by **B**

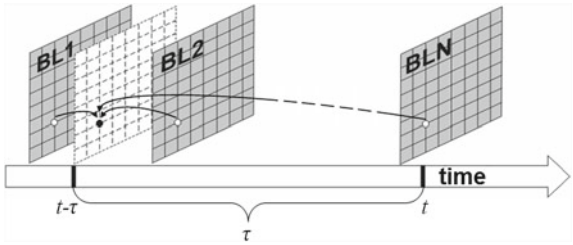


Fig. 2 Data interpolation scheme for non-base time slice (*white plane*) using data of base time slices stored in RAM

It is obvious that the main problem in the proposed method is restoring the nodal values for all non-base time slices. By taking into account that the number of base slices is time-dependent, we have used Newton's interpolation polynomial:

$$\begin{aligned} P_N(t-\tau,x,y,z) = {} & R(U_{BL1}(x,y,z)) + (t-\tau-t_{BL1}) \\ & R(U_{BL1}(x,y,z), U_{BL2}(x,y,z)) + \cdots \\ & + (t-\tau-t_{BL1})(t-\tau-t_{BL2})\ldots(t-\tau-t_{BLN}) \\ & R(U_{BL1}(x,y,z),\ldots,U_{BLN}(x,y,z)), \end{aligned} \tag{7}$$

where R are the divided differences of the corresponding order N, $U_{BLN}(x,y,z)$ and t_{BLN} are the function value and time for N-th base time slice respectively. The procedure (7) is performed for each grid node, where the values of base slices play the role of interpolation nodes. If the function **U** is sufficiently smooth, one can construct the polynomial using only two neighboring base layer, which significantly reduces the computation time (Fig. 2).

Finally, note that if the problem has a term depending on entire evolution of the system (history term), the technique described above does not change significantly. In this case one needs to keep the base time slices throughout the entire evolution of the system. Thus, the method can be easily generalized to more complex cases.

3 Numerical Examples

In this section we present three numerical examples of the application of the data storage optimization technique described above. We demonstrate that this procedure can be applied to the equations which include functional equations with a constant or variable time delay in one or more spatial dimensions, differential integral equations with a spatially non-local delay term and partial delay differential equations with complex spatio-temporal dynamics.

3.1 Partial Differential Equation with Constant Delay Admitting Exact Solution

The first example was taken from the paper by Zubik-Kowal and Vandewalle [11]. We consider a linear partial differential equation with a constant delay argument

$$\frac{\partial U(t,x)}{\partial t} = \frac{1}{10+40t^2}\frac{\partial^2 U(t,x)}{\partial x^2} + e^{-4t^2}U(t-1,x) + g(t,x), \tag{8}$$

for $t \in [0, 20]$, $x \in [-5, 5]$. The function $g(t, x)$ is selected in the form

$$g(t,x) = e^{-x^2}\left(1 + t\frac{1-2x^2}{5+20t^2} + (1-t)e^{-4t^2}\right) \tag{9}$$

in such a way that the exact solution of the problem (8, 9) becomes

$$U(t,x) = te^{-x^2}.$$

In fact, the equation (8, 9) can be classified as a delay partial differential equation in the one spatial dimension with a constant time delay. We define the numerical error as

$$E(t) = \max\left|U^{num}(t,ih) - U(t,ih)\right|. \tag{10}$$

The evolution of the error (10) as function of time for $t \in [0, 20]$ and $h = 0.05$ for different values of the parameter K is shown in Fig. 3. One can see that since the function is changing slowly, the adaptive method saves roughly each 20th time slice. In [11] the problem (8, 9) has been solved by the method of lines with the waveform relaxation technique. Comparing the accuracy of their method one can conclude that our adaptive method is approximately equal to Jacobi/Pickard waveform relaxation scheme with about 20 iterations at each time step.

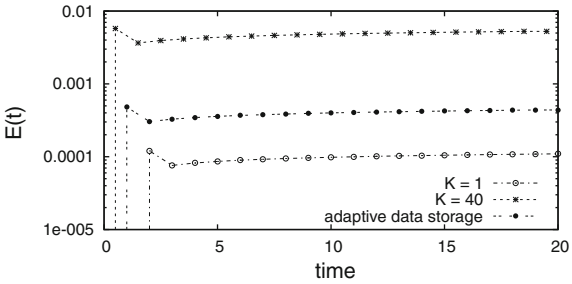


Fig. 3 Errors as a function of time for $t \in [0, 20]$ for Example #1 (8, 9)

3.2 Model of Spatial Spread of Infection: System of Reaction-Diffusion Equations Accounting for a Spatially Non-Local Term with a Discrete Delay

The second example is based on the model of spatial spread of infectious diseases with a fixed latent period suggested by Li and Zou [9]. It is assumed in the model that an infectious disease in a population has a fixed latent period and the latent individuals of the population may diffuse in the one spatial dimension. The model is given by the following system of non-linear reaction-diffusion equations with a discrete delay accounting for the latency and a spatially non-local term caused by the mobility of the individuals during the latent period:

$$\frac{\partial S(t,x)}{\partial t} = \mu + D_S \frac{\partial^2 S(t,x)}{\partial x^2} - dS(t,x) - rI(t,x)S(t,x), \tag{11}$$

$$\frac{\partial I(t,x)}{\partial t} = D_I \frac{\partial^2 I(t,x)}{\partial x^2} - \beta I(t,x) + \varepsilon \int_{-\infty}^{\infty} rI(t-\tau, y)S(t-\tau, y) f_\alpha(x-y)dy, \tag{12}$$

where $t > 0$, $x \in \Re$. Here $S(t,x)$ stands the susceptible part of the total population that consists of those individuals that can be infected; $I(t,x)$ denotes the individuals that capable of infecting others. In was found in [9] that the problem (11, 12) has the traveling wave solutions in some ranges of parameters. In that work the initial-boundary problem has been solved by the method of lines. However, authors have noted that the method of lines is very useful in solving a reaction diffusion system in the absence of time delay and spatial non-locality. But their model (11, 12) contains a time delay representing latent period as well as a non-local term resulted from the mobility of the latent individuals, and the method cannot be applied directly. So, they were forced to develop special technique in order to transform this system into one without non-locality and to apply the method of lines [9]. It is clear that the authors developed the technique for a concrete problem and this technique cannot be easily applied to other problems.

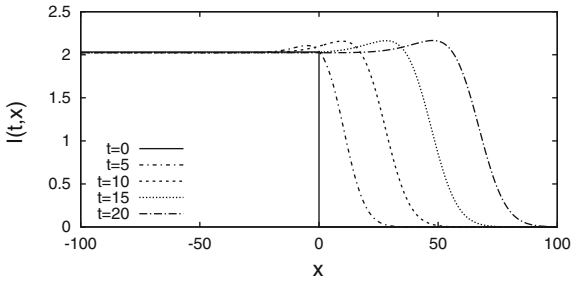


Fig. 4 Delay induced traveling wave observed in the system (11, 12) with parameters: $\mu = 5$, $d = 0.5$, $r = 0.5$, $D_S = 10$, $D_I = 10$, $\beta = 1$, $\tau = 1.005$

Our method is free from such disadvantages. It can be easily applied to the problem (11, 12). Figure 4 presents the traveling wave solution developing from step-function taken as the initial condition. It is found that the wave spreads at a speed $C = 4.96$ which is very close to the value $C = 4.9974$ obtained in [9].

3.3 Model of Spatially Extended Delay-Induced Circadian Rhythms

Finally, we show that our method can be applied to the problem in more than one spatial dimension. Let us consider the model of the circadian oscillations in *Neurospora crassa* developed by authors in [12]:

$$\frac{\partial F}{\partial t} = \frac{1}{1+4K_1^F F}\left(A_F + k_F \frac{K_1^W K_2^F W^2(t-\tau)}{1+K_1^W K_2^F W^2(t-\tau)} - B_F F - kFW\right) + D\left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}\right), \tag{13}$$

$$\frac{\partial W}{\partial t} = \frac{1}{1+4K_1^W W}\left(A_W + k_W \frac{K_1^F K_2^W F^2(t-\tau)}{1+K_1^F K_2^W F^2(t-\tau)} - B_W W - kFW\right) + D\left(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2}\right). \tag{14}$$

Here F and W stand for concentrations of proteins and D is the coefficient of protein diffusion in the cell. We define a two-dimensional domain Σ: $(0 < x < 200,$ $0 < y < 200)$ with zero-flux conditions for protein concentrations imposed at the boundary of the domain. The initial condition was a random distribution of the protein concentrations. In the very beginning of the simulation when one needs to know the spatial fields within the interval of delay, it was approximated by the harmonic

Table 1 The speedup of calculations using proposed algorithm of data storage. Each numerical integration was performed over range $0 < t < 1000$ starting from the same initial conditions

K	1	2	5	10	20	Adaptive algorithm
Time, *s*	7012	2320	913	720	665	909
Speedup, *times*	1	3.02	7.68	9.74	10.54	7.71

functions with random amplitude, phase and frequency. The initial-boundary value problem (13, 14) has been solved by finite difference method described in the previous section. The explicit scheme (3) was adopted to discretize equations. The equations and boundary conditions have been approximated on a rectangular uniform mesh 400×400 using a second order approximation for the spatial coordinates.

Our results have been tested to be robust with regard to time and space mesh refinements. All models were programmed in Delphi and simulated on AMD Athlon 64 3800+(2.41hHz) computer.

In the numerical calculations we have used the following fixed values of parameters: $K_1^{F,W} = 5$, $k_F = 8\,\text{nM/h}$, $k_W = 4$ nM/h, $K_2^{F,W} = 5$, $B_F = 0.3\ \text{h}^{-1}$, $B_W = 0.4\ \text{h}^{-1}$, $k = 30\ \text{nM}^{-1}\text{h}^{-1}$, $\tau = 6\,\text{h}$, $D = 0.01\ \text{m}^2\text{s}^{-1}$.

We have carried out the numerical experiments to see how the proposed method saves the computer resources (see Table 1). It is found that when $K = 2$ (each 2nd time slice is base), the calculation is about 3 times faster, but for $K = 20$ it is just 10.5 time faster. The reason for such nonlinear dependence is the growth of calculations related with the interpolation of the spatial fields.

The main question is how does the change of parameter K affect the accuracy of the resolution of spatio-temporal structures? Due to the strong nonlinearity of the system (13, 14), even small deviations of the period, amplitude and phase of spatially extended oscillations could lead to significant changes in the structure. Figure 5 presents the density plots of protein F concentration after 1000 h of evolution for the different numerical schemes. The first figure (Fig. 5a) corresponds to the standard approach when all the nodal data is stored in the computer memory ($K = 1$). If we store only half of the data ($K = 2$), the spatial distribution of reactant shown in Fig. 5b seems to be almost unchanged. In the case $K = 20$ the result of the evolution depicted in Fig. 5c changes qualitatively. It becomes evident when comparing the selected areas of the field with the reference field shown in Fig. 3a. For example, if $K = 20$ there is no any spiral wave in the selected area indicated in Fig. 5c. In contrast with it the spiral wave is clearly observed in Fig. 5a. The numerical scheme with the adaptive algorithm demonstrates an excellent result, since the final spatio-temporal structure shown in Fig. 5d coincides with the reference one. On average, the adaptive numerical scheme speeds up the calculation in 7.7 times.

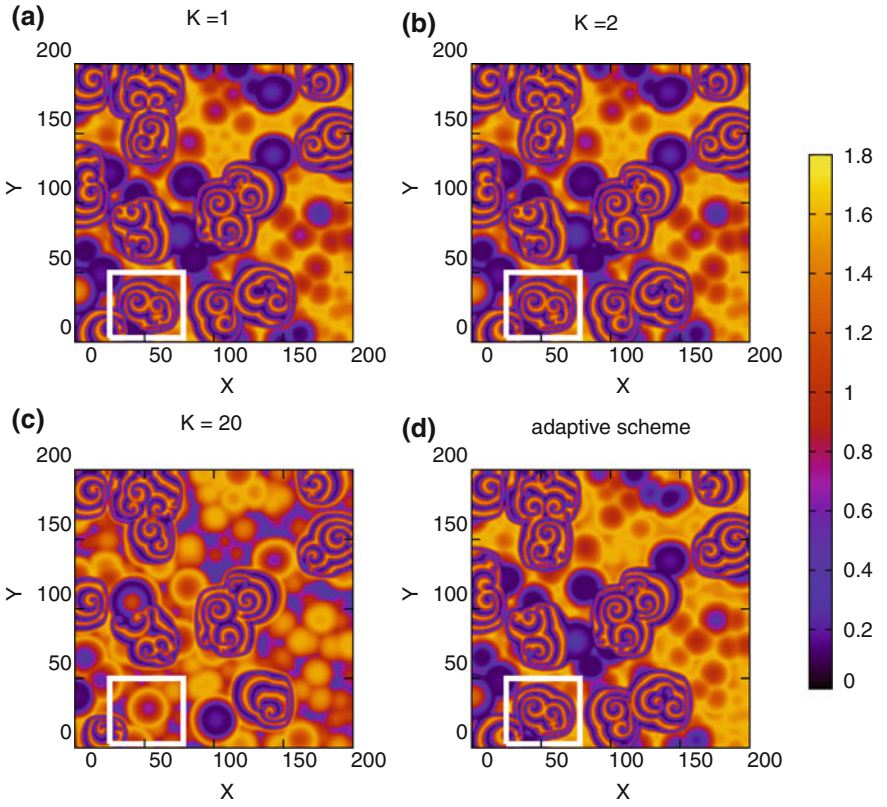


Fig. 5 Density plots of the concentration of the F protein at time $t = 1000$ for the different numerical schemes: **a** $K = 1$; **b** $K = 2$; **c** $K = 20$; **d** the adaptive algorithm

4 Conclusions

In order to study spatio-temporal dynamics of reaction-diffusion systems with time-delayed feedback, a novel adaptive algorithm of the numerical simulation of such systems has been proposed. The algorithm based on a finite difference method involves storing in a computer memory not all, but some selected nodal data, and the subsequent interpolation to determine intermediate values. The algorithm has been applied to simulate three different initial-boundary problems. It is shown that the adaptive numerical scheme can speed up the calculation in 7.7 times without visible changes in the amplitude and phase of spatially extended oscillations.

The work was supported by the Department of Science and Education of Perm region (project C26/244), the Ministry of Science and Education of Russia (project 1.3103.2011) and Perm State Pedagogical University (project 031-F).

References

1. Lakshmanan, M., Senthilkumar, D.V.: Dynamics of Nonlinear Time-Delay Systems. Springer, Berlin (2010)
2. Hale, J.K.: Theory of Functional Differential Equations. Springer, New York (1977)
3. Bratsun, D.A.: Effect of unsteady forces on the stability of non-isothermal particulate flow under finite-frequency vibrations. Microgravity Sci. Technol. **21**, 153–158 (2009)
4. Bratsun, D., Volfson, D., Hasty, J., Tsimring, L.S.: Delay-induced stochastic oscillations in gene regulation. Proc. Natl. Acad. Sci. U.S.A. **102**, 14593–14598 (2005)
5. Schiesser, W.E.: The Numerical Method of Lines: Integration of Partial Differential Equations. Academic Press, San Diego (1991)
6. Higham, D.J., Sardar, T.: Existence and stability of fixed points for a discretised nonlinear reaction-diffusion equation with delay. Appl. Numer. Math. **18**, 155–173 (1995)
7. Rey, A.D., Mackey, M.C.: Multistability and boundary layer development in a transport equation with delayed arguments. Canad. Appl. Math. Quart. **1**, 61–81 (1993)
8. Jackiewicz, Z., Zubik-Kowal, B.: Spectral collocation and waveform relaxation methods for nonlinear delay partial differential equations. Appl. Numer. Math. **56**, 433–443 (2006)
9. Li, J., Zou, X.: Modeling spatial spread of infectious diseases with a fixed latent period in a ppatially continuous domain. Bull. Math. Biol. **71**, 2048–2079 (2009)
10. Smolen, P., Baxter, D.A., Byrne, J.H.: Modeling circadian oscillations with interlocking positive and negative feedback loops. J. Neurosci. **21**, 6644–6656 (2001)
11. Zubik-Kowal, B., Vandewalle, S.: Waveform relaxation for functional-differential equations. SIAM J. Sci. Comput. **21**, 207–226 (1999)
12. Bratsun, D., Zakharov, A.: Deterministic modeling spatio-temporal dynamics of delay-induced circadian oscillations in *Neurospora crassa*. In: Sanayei, A., Zelinka, I., Rössler, O.E. (eds.) ISCS 2013. Emergence, Complexity and Computation, Vol. 8, pp. 1-11. Springer, Heidelberg (2014)

Extracting the QRS Complexity and R Beats in Electrocardiogram Signals Using the Hilbert Transform

Ricardo Rodríguez, Adriana Mexicano, Salvador Cervantes, Jiri Bila and Rafael Ponce

Abstract This paper presents a novel approach for the problem of detecting and extracting the QRS complex of electrocardiogram signals for different kinds of arrhythmias. First, an autocorrelation function is used in order to obtain the period of an electrocardiagram signal and then the Hilbert transform is applied to obtain R-peaks and beats. Twenty three different records extracted from the MIT-BIH arrhythmia database were used to validate the proposed approach. In this testing has been observed a 99.9 % of accuracy in detecting the QRS complexity, being a positive result in comparison with other recent researches.

Keywords Hilbert transform · Electrocardiogram signals · Autocorrelation

1 Introduction

The processing of Electrocardiogram (ECG) signals has become an important area in medicine [1]. It helps to detect and predict the majority of heart ailments [2]. A medical doctor is able to interpret ECG signals according to the knowledge and

R. Rodríguez (✉)
Technological University of Ciudad Juarez, Chihuahua, Mexico
e-mail: ricardo_rodriguez@utcj.edu.mx

A. Mexicano · S. Cervantes · R. Ponce
Polytechnic University of the State of Morelos, Morelos, Mexico
e-mail: amexicano@upemor.edu.mx

S. Cervantes
e-mail: scervantes@upemor.edu.mx

R. Ponce
e-mail: rponce@upemor.edu.mx

J. Bila
Czech Technical University in Prague, Prague, Czech Republic
e-mail: Bila@vc.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_20,

experiences obtained overtime. However, some human mistakes can appear while interpreting the signal, mainly due to recorded vibrations or the scale of the waveform [2, 3]. ECG records are obtained by sampling bioelectric currents and are composed of a *P* wave, a *QRS* complex and a *T* wave; the *QRS* represents ventricular depolarization and is the most important part of the ECG signal [4], from this is obtained the *R* wave and the maximum amplitude in the *R* wave is known as the *R*-peak. ECG processing presents many challenges for signal analysis and pattern recognition because the ECG signal extracting process can produce unwanted changes and variations in the signal recording and reduce the reliability of some sensitive parameters of *QRS* complex [5, 6]. Goya et al. in [2] claimed that it is necessary the use of an adequate signal processing to yield reliable diagnostic parameters for clinical diagnostics. In this sense, several signal processing algorithms [4, 7–18] have been developed to noise removal, data compression and parameter extraction [19].

For example in [4] an EMD algorithm (empirical mode decomposition method) has been used for QRS complex detection; similarly, in [7] an ensemble empirical mode decomposition (EEMD) and the K-nearest neighbors (K-NN) classifier were used. In [8] classification with support vector machine were used for ventricular fibrillation detection on ECG signals. In [9] to improve the classification, different wides were applied to remove *P* and *T* waves using the spectral entropy, heartbeat interval features and the *RR*-interval. In [10] the Hilbert transform is reported to be used for R-peak and QRS onset and offset detection, in [11] is used the first differential of the ECG signal and its Hilbert transformed data to locate R wave peaks.

In [12] wavelets were used to reduce the training time and to denoise the ECG signal in combination with Savitsky-Golay filter [13] or with a low-pass-filter [14]. In [15] the use of autocorrelation function is reported to improve the noisy immunity on heart rate estimation over fetal electrocardiography (FECG). In [16] and [17] adaptive low-pass filters and Hilbert transform respectively were used to remove noise peaks from the ECG signal. In [18] powerline interference was reduced with a notch filter and high frequency noise was suppressed with a 40 Hz lowpass filter.

This chapter is focused on the analysis of ECG signals by means of applying the autocorrelation function to obtain the period of a cardiac cycle and the application of the Hilbert transform to detect the real *R*-peaks from an ECG. Obtained results shows that the performance of the proposed method reported a sensitivity of 99.94 % and a 99.73 % of accuracy in QRS complex detection.

The rest of the paper is organized as follows: Sect. 2 presents a brief description of electrocardiograms signals, the band pass filter, the autocorrelation function and the Hilbert transform used in this work for detecting the QRS complex from an ECG. Section 3 shows the obtained results after applying the proposed method. Finally, Sect. 4 presents conclusions and future work.

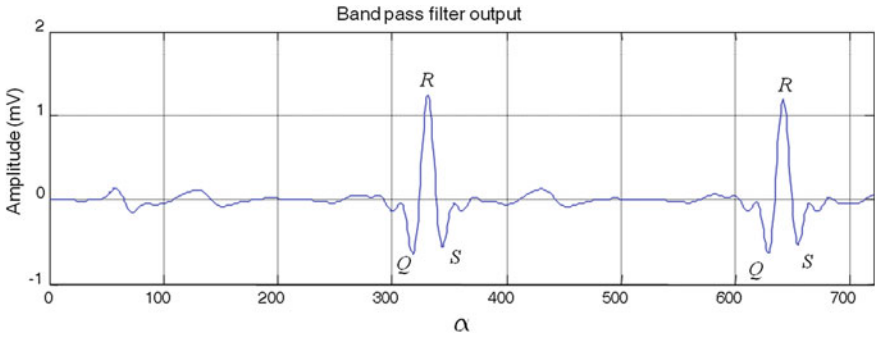


Fig. 1 Band pass filter output of the ECG signal of record 103 in MIT-BIH database [20]

2 Extracting the QRS Complexity Using Autocorrelation and the Hilbert Transform

ECG signals allow to represent the cyclical contraction and relaxation of human heart muscles. Heart muscle activity is controlled by electrical pulses which are transmitted through a nerve network; such electrical pulses are enough strong to be sensed by electrodes placed on the human skin [2, 21]. In general, the complex of an ECG signal is represented by the *QRS* variables as shown in Fig. 1. In order to emphasize the QRS complex, a band pass filter with pass band frequencies can be used. This filter is used to remove high frequency and power-line noise from the original signal. The band pass filter is given by the convolution of the original signal with the low-pass filter and high-pass filter. In time-domain the pass-band filter is given by (1)

$$\mathbf{r}(\alpha) = ((\mathbf{x}(\alpha) * \mathbf{h}_1(\alpha)) * \mathbf{h}_2(\alpha)) \tag{1}$$

where $\mathbf{x}(\alpha)$ stands for the raw ECG signal, $\mathbf{h}_1(\alpha)$ is the low pass filter, $\mathbf{h}_2(\alpha)$ is the high pass filter, and $\mathbf{r}(\alpha)$ is the band pass filtered of the ECG output. In this work, a band pass filter of (5–15) Hz was used; the band pass filter output of the channel-1 is presented in Fig. 1 The derivative output of the ECG signal allows to remove baseline drifts and motion artifacts. The derivative of the filtered ECG signal was calculated to obtain the fall from the R peak to S (see Fig. 2 which indicates the minimum slope of the filtered ECG signal). The derivative output also presents the high slope points which indicate that the rise of the signal from Q to R corresponds to the maximum slope [22]. The first order differentiation of the filtered ECG can be obtained using (2)

$$\mathbf{r}(\alpha) = \mathbf{r}(\alpha) - \mathbf{r}(\alpha - 1), \alpha = 0, 1, \ldots, n - 1 \tag{2}$$

where n is the total number of samples. Figure 2 shows the first order differentiation of the filtered ECG. The autocorrelation function of the ECG time series is calculated by the analysis of the filtered ECG signal dynamics. Autocorrelation is used to detect

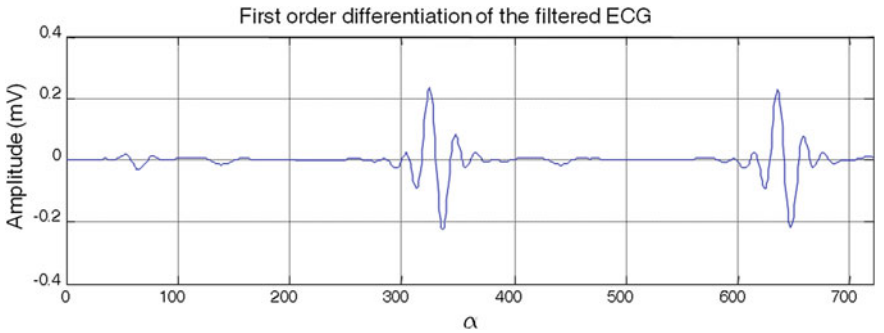


Fig. 2 Derivative output of ECG filtered signal of record 103 in MIT-BIH database, window of 2 s length

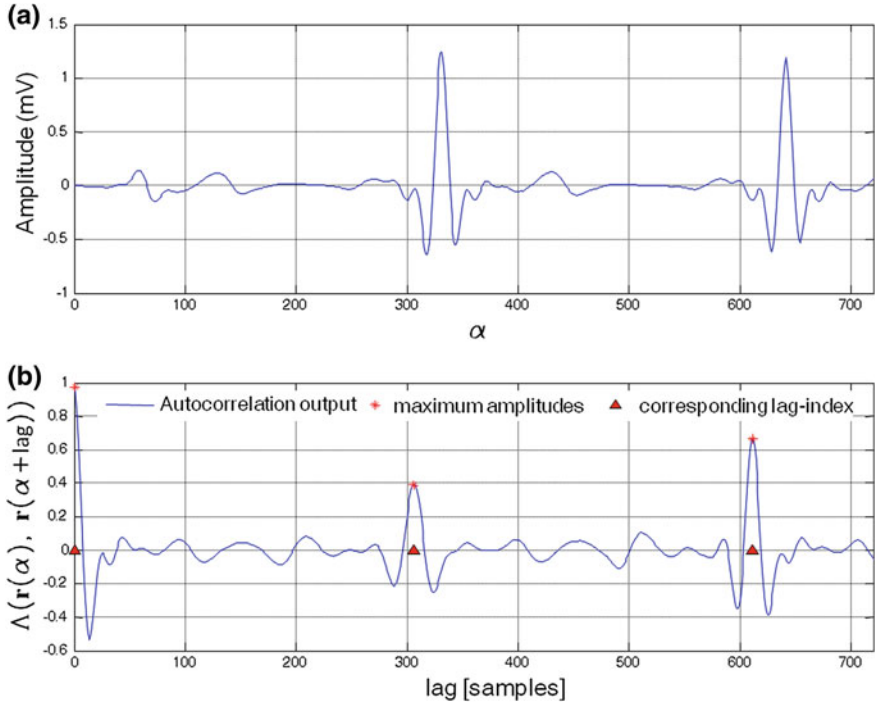


Fig. 3 Autocorrelation function of the filtered ECG signal dynamics for one instance of a window of 2 s = 720 steps. **a** represents the filtered ECG signal. **b** represents the autocorrelation function of the above filtered ECG signal

the similarities between the signals $\mathbf{r}(\alpha)$ and $\mathbf{r}(\alpha + h)$ with a time-lag of factor h. In the proposed calculation a duration of 2 s for the filtered ECG is considered in order to find the exact duration of one cardiac cycle (for this signal). The autocorrelation function used in this work is calculated by (3)

$$\Lambda_h = \frac{\mathcal{E}\left[(\mathbf{r}_\alpha - \mu_\alpha)(\mathbf{r}_{\alpha+h} - \mu_{\alpha+h})\right]}{\sigma_\alpha \sigma_{\alpha+h}} \tag{3}$$

where the time-lag of $h = 1, \ldots, n$. Figure 3 shows a 2 s window of the filtered ECG signal from the MIT-BIT record 103. In Fig. 3a, the filtered ECG signal is shown. In Fig. 3b, autocorrelation function of the filtered signal is presented. The approximate interval between two cardiac cycles according to the autocorrelation function is around 290 samples, it means around 0.8 s. According to [22] the approximate *R-R* interval between two cardiac cycles is from 0.4 to 1.2 s. In Fig. 3b, it is shown that the maximum amplitude shows a three cycle duration of the filtered signal, which was obtained from the output of the autocorrelation. For the time varying ECG signal, the Hilbert transform is used for envelope detection. The detector is based on the envelope of the Hilbert transform as can be seen in Fig. 3. In Fig. 3a, the envelope of the Hilbert transform is shown superimposed alongside the differentiation output. Figure 3b, shows the Hilbert transform of the filtered ECG time series. The maximum peak obtained in the envelope of the Hilbert transform output is the zero crossing point of the differentiation output. The Hilbert transform of $\mathbf{r}(\alpha)$ is given by (4).

$$\widehat{\mathbf{r}}(\alpha) = \mathcal{H}(\alpha) = \mathcal{FFT}^{-1}(\mathbf{x}(\alpha) * \mathbf{h}(\alpha)) \tag{4}$$

where $\mathcal{FFT}^{-1}$ stands for the Inverse Fast Fourier Transform, the vector $\mathbf{h}$ is created with the values in (5)

$$\begin{array}{l} 1 \text{ for } i = 1, n/2 + 1 \\ 2 \text{ for } i = 2, 3, \ldots, n/2 \\ 0 \text{ for } i = n/2 + 2, \ldots, n \end{array} \tag{5}$$

and the vector $\mathbf{x}$ stores the Fast Fourier Transform ($\mathcal{FFT}$) of the ECG signal $\mathbf{r}(\alpha)$. The $\mathcal{FFT}$ is given by (6)

$$\mathbf{x}(\alpha) = \sum_{i=1}^{n} \mathbf{r}(i) \bullet \omega_n^{(i-1)(\alpha-1)} \tag{6}$$

where n is the length of ECG signal $\mathbf{r}(i)$, and $\omega_n = e^{(-2\pi i)/n}$ is an nth root of unity.

The Hilbert transform, as shown in (4), is obtained by solving the inverse $\mathcal{FFT}$ of the convolution of the vectors $\mathbf{x}(\alpha)$ and $\mathbf{h}(\alpha)$ using (7)

$$\mathcal{H}(\alpha) = (1/n) \sum_{\alpha=1}^{n} (\mathbf{x}(\alpha) * \mathbf{h}(\alpha)) \omega_n^{-(i-1)(\alpha-1)} \tag{7}$$

The analytic signal is expressed as in (8)

$$r_a = r(\alpha) + i\widehat{r}(\alpha) \tag{8}$$

And its instantaneous phase angle in the complex plane can be defined as (9)

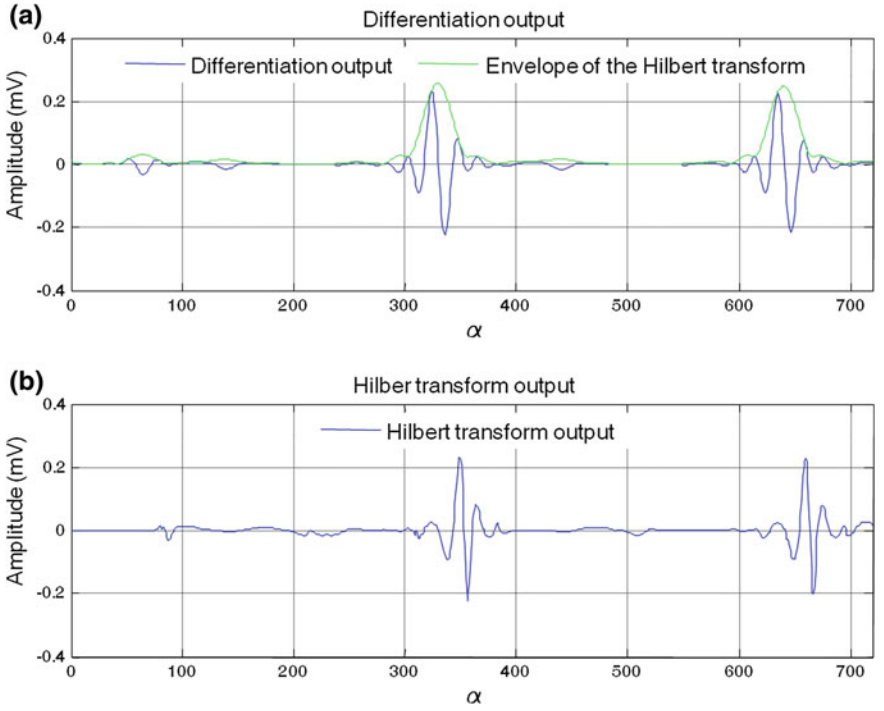


Fig. 4 Hilbert transform of the differentiation output of the ECG signal dynamics for one instance of a window of 720 samples (i.e., 2 s). **a** shows the differentiation output of the ECG signal, and the envelope of the Hilbert transform superimposed on its differentiation output. **b** the Hilbert transform of the above ECG signal is shown

$$\Theta(\alpha) = tan^{-1}\left(\frac{\widehat{r}(\alpha)}{r(\alpha)}\right) \tag{9}$$

The envelope $\mathcal{B}$ of $\Theta(\alpha)$ is presented by (10)

$$\mathcal{B} = \sqrt{r^2(\alpha) + \widehat{r}^2(\alpha)} \tag{10}$$

The envelope $\mathcal{B}$ is always positive, and it will have the same magnitude and slope of the original signal $\mathbf{r}(\alpha)$ at or in the vicinity of its local maxima.

The maximum point of the Hilbert transform envelope represents the *R*-peak point in the QRS signal. However, it is an approximation of the current R-peak of the ECG signal (see Fig. 4). To identify the real *R*-peak point, the maximum point is searched within a fixed window of 74 samples from the position of the *R*-peak.

After that, the *Q* point is located by means of a search of the minimum point in a window of -13 samples back from the *R*-peak position. In a similar way, the *S* point is located by searching the minimum inside a window of +13 samples from

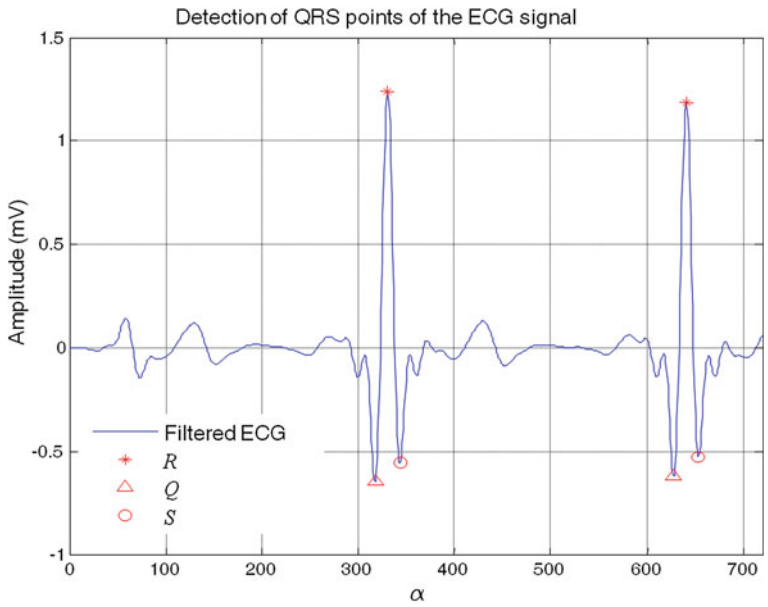


Fig. 5 Detection of the *QRS* points of filtered ECG signal

the *R*-peak position. Figure 5 shows the detected QRS points in the ECG signal using the methodology presented in this paper. Once we obtain the QRS points of the filtered ECG signal we determine the number of beats per minute of the ECG signal. In order to perform this, the number of R-peaks detected by the Hilbert transform were counted. Then, the duration in seconds was obtained by dividing the duration in minutes over 60. Finally, the number of beats per minute (bpm) was obtained by dividing the number of R-peaks (beatcounts) over the duration in minutes, as can be seen in (11)

$$bpm = \frac{beatcounts}{duration\_in\_minutes} \tag{11}$$

3 Computational Results

In order to validate our approach, 23 different records (100–124) of the MIT-BIH arrhythmia database [20] were used. The records were sampled at 360 Hz, with 11-bit resolution over ±5 mV range. Each record has 30 min and 5.556 s of two channel long. The performance was analyzed using the following parameters:

a) Sensitivity: *Se*(%) calculated as shown in (12) indicates the percentage of correctly classified beats among all beats.

$$Se(\%) = \frac{TP}{TP + FN} * 100 \tag{12}$$

b) Positive predictivity: $+P(\%)$ calculated as shown in (13) indicates the percentage of the correctly detected beats from the total number of beats.

$$+P(\%) = \frac{TP}{TP + FP} * 100 \tag{13}$$

c) Detection error rate: $e(\%)$ calculated as shown in (14) corresponds to the percentage of false detections over the total number of detected heartbeats.

$$e(\%) = \frac{FP + FN}{TP + FN} * 100 \tag{14}$$

where, True Positives(TP) represent the number of positive detections that corresponds to the annotations of the specialist, False Positives(FP) stand for the number of detections that do not correspond with the annotations of the specialist and False Negatives (FN) correspond to the number of heart beats that were annotated by the specialists, that were not detected. The performance of our method using the MIT-BIH database is shown in Table 1. On average, the proposed method for detecting the QRS complex presents a 0.10 % of detection error rate, a positive prediction of 99.73 % and a sensitivity of 99.94 %.

In comparison with other similar researches, [4] using an EMD algorithm reported a positive predictivity of 99.97 % and sensitivity of 99.97 % for the *R*-Peak detection, using records from Fantasia database. Benitez [11] using the Hilbert transform on the MIT-BIH database had a positive prediction of 99.93 %, a detection error rate less than 0.8 % and sensitivity of 99.94 % meanwhile in [23] is reported a sensitivity of 99.95 %, a positive predictivity of 99.94 % and a detection error rate of 0.12 % (see Table 2)

From this comparison, it can be observed that the proposed approach using autocorrelation function and Hilbert transform produces competitive results against other approaches, and in this case, with a better detection error rate. Future testing of this approach with Fantasia database could compare the results obtained against the reported ones in [4].

4 Conclusions

This paper has presented a method for detecting the *QRS* complex of ECG signal, using autocorrelation function and the Hilbert transform techniques. Results show that applying our method we obtain an average positive predictive value of 99.73 %, Sensitivity of 99.94 % and detection error rate of 0.10 % in the *QRS* complex for the (100–124) records of MIT-BIT arrhythmia database. Because of the complicated

Table 1 Results of the proposed method in terms of Sensitivity, Positive predictive, and Detection error rate for the records (100–124) in MIT-BIH arrhythmia database

MIT-BIH record	$Se(\%)$	$+P(\%)$	$e(\%)$
100	100.00	100.00	0.00
101	100.00	94.95	0.05
102	99.86	100.00	0.13
103	100.00	100.00	0.00
104	99.82	99.77	0.44
105	99.76	99.76	0.46
106	99.90	99.95	0.15
107	99.95	99.85	0.18
108	99.88	100.00	0.11
109	100.00	99.80	0.19
111	99.90	99.95	0.14
112	100.00	100.00	0.00
113	100.00	100.00	0.00
114	99.80	99.78	0.31
115	100.00	100.00	0.00
116	99.95	100.00	0.04
117	99.93	100.00	0.06
118	100.00	100.00	0.00
119	100.00	99.89	0.10
121	99.94	100.00	0.05
122	100.00	100.00	0.00
123	100.00	100.00	0.00
124	100.00	100.00	0.00
Average	99.94	99.73	0.10

Table 2 Comparative of the proposed work against related work

Work	$Se(\%)$	$+P(\%)$	e(%)
Proposed	99.94	99.73	0.10
[11]	99.94	99.93	0.80
[23]	99.95	99.94	0.12

dynamics of the ECG signals, and its time-variant nature, we consider that, obtained results are promising for extracting the QRS complexity and R beats in electrocardiogram signals.

As a future work we propose the feature extraction of each heart beat and automate the classification of cardiac arrhythmias using multilayer perceptron neural network with backpropagation learning technique.

References

1. Karsikas, M.: New methods for vectorcardiographic signal processing, Acta universitatis Oululensis, University of Oulu. PhD thesis, Oulu (2011)
2. Goya-Esteban, R., Barquero-Perez, O., Alonso-Atienza, F., Ervess, E., Requena-Carrion, J., Garcia-Albeola, A., Rojo-Alvarez, J.L.: A review on recent patents in digital processing for cardiac electric signals (I): from basic systems to arrhythmia analysis. Recent Pat. Biomed. Eng. **2**, 22–31 (2009)
3. Asirvadam, V.S., Pisal, K.S., Izhar, L.I., Khuzi, N.A.A.M.: ECG viewed using grayscale patterns. In: Proceedings of the International Conference on Man-Machine Systems, pp. 11–13, Malaysia (2009)
4. Bagde, S., Raikwar, P.: Detection of QRS complexes of ECG waveform based on empirical mode decomposition using MATLAB. Inte. J. Eng. Innovative Technol. **1**(1), 14–17 (2012)
5. Karsikas, M., Huikuri, H., Perkiömäki, J.S., Lehtola, L., Seppänen, T.: Influence of paper electrocardiogram digitizing on T wave and QRS complex morphology parameters. Ann. Noninvasive Electrocardiol. **12l**, 282–290 (2007)
6. Lehtola, L., Karsikas, M., Koskinen, M., Huikuri, H., Seppänen, T.: Effects of noise and filtering on SVD-based morphological parameters of the T wave in the ECG. J. Med. Eng. Technol. **32**, 400–407 (2008)
7. Zhao, Z., Yang, L., Chen, D., Luo, Y.: A human ECG identification system based on ensemble empirical mode decomposition. Sensors **13**, 6832–6864 (2013)
8. Shandilya, S., Ward, K., Kurz, M., Najarian, K.: Non-linear dynamical signal characterization for prediction of defibrillation success through machine learning. BMC Med. Inform. Decis. Making **12**(116), 1–9 (2012)
9. Niwas, I., Selva, S., Sadasivam, V.: Artificial neural network based automatic cardiac abnormalities classification. In: Proceedings of the Sixth International Conference on Computational Intelligence and Multimedia Applications, pp. 41–46 (2009)
10. Illanes, A., Zhang, Q.: An algorithm for QRS onset and offset detection in single lead electrocardiogram records. In: Proceedings of the 29th Annual International Conference of the IEEE EMBS Cité Internationale, pp. 541–544 (2007)
11. Benitez, D., Gaydecki, P., Zaidi, A., Fitzpatrick, A.: The use of the hilbert transform in ECG signal analysis. Comput. Biol. Med. **31**, 399–406 (2001)
12. Özbay, Y.: A new approach to detection of ECG arrhythmias: complex discrete wavelet transform based complex valued artificial neural network. J. Med. Syst. **33**, 435–445 (2009)
13. Ebrahimzadeh, A., Khazaee, A.: Detection of premature ventricular contractions using MLP neural networks: a comparative study. Measurement **43**(1), 103–112 (2010)
14. Monasterio, V., Laguna, P., Martnez, J.: Multilead analysis of T-Wave alternans in the ECG using principal component analysis. IEEE Trans. Biomed. Eng. **56**, 1880–1890 (2009)
15. Jezewski, J., Roj, D., Wrobel, J., Horoba, K.: A novel technique for fetal heart rate estimation from doppler ultrasound signal. Biomed. Eng. Online **10**, 92–92 (2011)
16. Zong, C., Chetouani, M.: Hilbert-Huang transform based physiological signals analysis for emotion recognition. In: Proceedings of the IEEE International Symposium on Signal Processing and Information Technology, pp. 334–339 (2009)
17. Kohli, S., Makwana, N., Mishra, N., Sagar, B.: Hilbert transform based adaptive ECG R-Peak detection technique. Int. J. Electr. Comput. Eng. **2**(5), 639–643 (2012)
18. Kentta, T., Karsikas, M., Juhani, M., Juha, S., Seppanen, T., Kiviniemi, A., Nieminen, T., Lehtimaki, T., Nikus, K., Lehtinen, R., Viik, J., Kahonen, M., Huikuri, H.: QRS-T morphology measured from exercise electrocardiogram as a predictor of cardiac mortality. Europace **13**, 701–707 (2011)
19. Acar, B., Yi, G., Hnatkova, K.: Spatial, temporal and wavefront direction characteristics of 12-lead T-wave morphology. Med. biol. Eng. comput. **37**, 574–584 (1999)
20. MIT-BIH Database distribution, http://www.physionet.org/physiobank/database/mitdb

21. Kotas, M.: Projective filtering of time-aligned ECG beats for repolarization duration measurement. Comput. Methods Program Biomed. **85**(2), 115–123 (2007)
22. Yeh, Y.C., Wang, W.J.: QRS complex detection for ECG signal: the difference operation method. Comput. Methods Program Biomed. **91**(3), 245–254 (2008)
23. Prakash, J.: Analysis of ECG signal for Detection of Cardiac Arrhythmias. National Institute Of Technology, Master of technology thesis, Rourkela (2011)

Analyses of the Chaotic Behavior of the Electricity Price Series

Radko Kříž and Štěpán Kratochvíl

Abstract Electricity price is by its features like mean-reversion, high volatility rate and frequent occurrence of jumps different from other commodities. These differences are mainly caused by non-storability of the electricity, which need to balance supply and demand in real time. Due to these features, electricity price behavior seems somewhat chaotic. In this chapter we will introduce methods for investigating whether or not electricity spot prices can be described by usual time series, stochastic models. For this reason we will estimate the largest Lyapunov exponent and the 0–1 test for chaos. We will do a case study on the EPEX Phelix spot index.

Keywords Chaos theory · Electricity spot price · Time series analysis · Phase space reconstruction · Hurst exponent · Lyapunov exponents · 0–1 test for chaos

1 Introduction

Forecasting electricity energy prices began in the 1990s, at the beginning of the liberalization process. It is very important for electricity traders to forecast electricity prices with the highest possible precision as the precision of the forecast is connected to the trading strategy and thus with profit or lost.

The price depends on the demand and supply and on the costs of transportation in the grid. The first two features are strongly dependent on the weather, economic situation, government interventions etc. These dependencies result in a very complex fluctuation of electricity prices, which may seem to be slightly chaotic. The price is

R. Kříž (✉) · Š. Kratochvíl
Department of Economics, Management and Humanities, Faculty of Electrical Engineering, Czech Technical University, Technická 2, 166 27 Praha, Czech Republic
e-mail: krizradk@fel.cvut.cz

Š. Kratochvíl
e-mail: kratoste@fel.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_21,

set by an interaction between supply and demand and in set in auction at a single time for the whole 24 h of the following day. This approach is unlike the usage of time series models (autoregressive type models like AR, ARMA or ARIMA) as it goes against the assumption, that with another value (another hour), the information set would be extended.

The electricity spot prices can also be split up into deterministic and stochastic parts. The deterministic part describes the seasonality and all the periodic patterns (intra-day, intra-weak, peak, weekend patterns). This part is easily observable and we can describe it simply. The stochastic part (also called volatility), on the contrary, is unobservable and seems chaotic. GARCH type models can be used in applying the stochastic approach for modeling volatility. Unfortunately these models can be applied untill chaotic behavior is proved.

As was noted above, if the need for chaotic description is proven, it will not be possible to use the usual stochastic and time series methods such as the autoregressive and stochastic models and the regime switching model. In this case, the methods which were developed specially to describe time series with chaotic behavior should be used.

There are some techniques currently used for forecasting electricity prices under chaotic description. The most frequently used is the NN (Neural Network) which detects patterns in the previous samples and is able to forecast new samples thanks to the Learning Phase Geva [1]. On the basis of NN, the RNN (Recurrent Neural Network), which is an autoregressive nonlinear dynamic model capable of representing arbitrary nonlinear dynamic systems, was developed. This model is unique with its feedback loop, which makes it possible to retain a memory of previous states Mirikitani and Nikolaev [2].

Another approach is to use wavelet transformation, where signals are decomposed and then reconstructed in the given phase space. Reconstruction based on the Phase Space is the basic method for forecasting new data. However, with a higher embedding dimension (see below), it becomes very difficult to obtain a description. The graphical solution can be found when the embedding dimension equals two or three, but with a higher embedding dimension numerical analysis has to be done.

To be able to use these methods and models we will have to prove first that electricity spot prices behave chaotically and that it is adequate to use a chaotic description. The basic question is therefore the existence of chaotic behavior. If the system behaves chaotically, we are forced to accept only limited predictions. Nevertheless chaotic behavior is much better than random processes.

Thus we will focus in this chapter on the existence of chaotic behavior in the hourly EPEX Phelix spot index.

This chapter is organized as follows. Section 2 presents the theoretical introduction and method for computing. The Analysis of electricity price series is in Sect. 3. Input data are described in Sect. 3.1. The following subsections show the calculation of time delay, embedding dimension, the largest Lyapunov exponent, 0–1 Test, Hurst exponent and fractal dimensions. The conclusion is in Sect. 4.

2 Analysis Methodology

In short, we will describe the basic definitions and the basic methods for examining the input data.

2.1 Phase Space Reconstruction

For the description of the data under chaotic behavior (nonlinear time series), we need to estimate phase space, which specifies whether or not a given time series is of a deterministic nature. For construction of the phase space we need to define the dimension of the data and the time delay, which specifies the size of phase space [3].

The key to answering the question of whether the data set is chaotic, is embodied in the method of phase space reconstruction, which has been rigorously proven by the embedding theorems of Takens [4]. Takens' theorem was independently suggested for example by Packard [5]. Takens' theorem transforms the prediction problem from time extrapolation to phase space interpolation.

Let us take a given time series $x_1, x_2, \ldots, x_N$ which is embedded into the m-dimensional phase space by the time delay vectors. A point in the phase space is given as:

$$Y_n = x_n, x_{n-\tau}, \ldots, x_{n-(m-1)\tau} \quad n = 1, 2, ..., N - (m-1)\tau \tag{1}$$

where τ is the time delay and m is the embedding dimension. Different choices of τ and m yield different reconstructed trajectories. How can we determine optimal τ and m?

2.2 Optimal Time Delay

A one-to-one embedding can be obtained for any positive value of the time delay $\tau > 0$. However, very small time delays will result in near-linear reconstructions with high correlations between consecutive phase space points, and very large delays might obscure the deterministic structure linking points along a single degree of freedom [6].

In order to estimate τ, two criteria are important according to Kodba [7]. First, τ has to be large enough so that the information we get from measuring the value of x at time $n+\tau$ is significantly different from the information we already have by knowing the value of x at time n. Only then will it be possible to gather enough information about all other system variables that influence the value of x to reconstruct the whole attractor. Second, τ should not be larger than the typical time in which the system loses memory of its initial state. This is particularly important for chaotic systems,

which are intrinsically unpredictable and hence lose memory of the initial state as time progresses.

Following this reasoning, Fraser and Swinney [8] introduced the mutual information between x_n and $x_{n+\tau}$ as a suitable quantity for determining τ. The mutual information between x_n and $x_{n+\tau}$ quantifies the amount of information we have about the state $x_{n+\tau}$ presuming we know the state x_n. Now we can define the mutual information function:

$$I(\tau) = -\sum_{h=1}^{j}\sum_{k=1}^{j} P_{h,k}(\tau) \ln \frac{P_{h,k}(\tau)}{P_h P_k} \tag{2}$$

where P_h and P_k denote the probabilities that the variable assumes a value inside the hth and kth bins, respectively, and $P_{h,k}(\tau)$ is the joint probability that $\mathrm{x_n}$ is in bin h and $x_{n+\tau}$ is in bin k. Hence, the first minimum of $I(\tau)$ marks the optimal choice for the time delay.

2.3 Optimal Embedding Dimension

The embedding dimension m is conventionally chosen using the "false nearest neighbors" method. The minimum embedding dimension capable of containing the reconstructed attractor is that for which the percentage of false nearest neighbors drops to zero for a given tolerance level ε [6].

In order to calculate the fraction of false nearest neighbors, the following algorithm is used according to Kennel [9]. Given a point $p(i)$ in the m-dimensional embedding space, one first has to find a neighbour $p(j)$, so that:

$$\|p(i) - p(j)\| \leq \varepsilon \tag{3}$$

We then calculate the normalized distance R_i between the $(m+1)$th embedding coordinate of points $p(i)$ and $p(j)$ according to the equation:

$$R_i = \frac{\left|x_{i+m\tau} - x_{j+m\tau}\right|}{\|p(i) - p(j)\|} \tag{4}$$

If R_i is larger than a given threshold R_{tr}, then $p(i)$ is marked as having a false nearest neighbor. Equation (4) has to be applied for the whole time series and for various $m = 1, 2, \ldots$ until the fraction of points for which $R_i > R_{tr}$ is negligible [7].

2.4 Lyapunov Exponents

Lyapunov exponent or Lyapunov characteristic exponent of a dynamical system is a quantity that characterizes the rate of separation of infinitesimally close trajectories. Quantitatively, two trajectories in phase space with initial separation δZ_0 diverge (provided that the divergence can be treated within the linearized approximation) [10].

$$\delta Z(t) \approx e^{\lambda t} \, |\delta Z_0| \tag{5}$$

where λ is the Lyapunov exponent.

The largest Lyapunov exponent can be defined as follows:

$$\lambda = \lim_{\substack{\delta Z_0 \to 0 \\ t \to \infty}} \frac{1}{t} \ln \frac{|\delta Z(t)|}{|\delta Z_0|} \tag{6}$$

The limit $\delta Z_0 \to 0$ ensures the validity of the linear approximation at any time.

Largest Lyapunov exponent determines a notion of predictability for a dynamical system. A positive largest Lyapunov exponent is usually taken as an indication that the system is chaotic (provided some other conditions are met, e.g., phase space compactness).

In our work we have used the Rosenstein algorithm, which calculates the largest Lyapunov exponent as follows:

$$\lambda_1(i) = \frac{1}{i\Delta t} \cdot \frac{1}{(M-i)} \sum_{j=1}^{M-i} \ln \frac{d_j(i)}{d_j(0)} \tag{7}$$

where $d_j(i)$ is distance from the j point to its nearest neighbor after i time steps and M is the number of reconstructed points. For more information see [11, 12].

2.5 The 0–1 Test for Chaos

A new test for the presence of deterministic chaos was developed by Gottwald and Melbourne [13]. Their '0–1' test for chaos takes as input a time series of measurements, and returns a single scalar value usually in the range 0–1. The 0–1 test does not depend on phase space reconstruction but rather works directly with the time series given. The input is the time-series data and the output is 0 or 1, depending on whether the dynamics are non-chaotic or chaotic [6].

Briefly, the 0–1 test takes as input a scalar time series of observations $\varphi_1, \ldots, \varphi_N$. We have used the algorithm according to Dawes and Freeland [14]. First, we must fix a real parameter c and construct the Fourier transformed series:

$$z_n = \sum_{j=1}^{n} \phi_j e^{ijc}, \; n = 1, \ldots, N \tag{8}$$

Then we have computed the smoothed mean square displacement:

$$M_c(n) = \frac{1}{N-p} \sum_{j=1}^{N-p} \left| z_{j+n} - z_j \right|^2 - \left(\sum_{k=1}^{N} \frac{\phi_k}{N} \right)^2 \frac{1 - \cos nc}{1 - \cos c} \tag{9}$$

Finally, we have estimated the correlation coefficient to evaluate the strength of the linear growth

$$r_c = \frac{\operatorname{cov}(n, M_c(n))}{\sqrt{\operatorname{cov}(n, n) \operatorname{cov}(M_c(n), M_c(n))}} \tag{10}$$

2.6 Hurst Exponent

Real processes in nature, according to the expectation of Benoit Mandelbrot, lie somewhere between pure deterministic process and white noise. This is why we can describe reality either by a stochastic or deterministic model. The Hurst coefficient can give us an answer to this. The values of the Hurst exponent vary between 0 and 1, with higher values indicating a smoother trend, less volatility, and less roughness. Random walk has a Hurst exponent of 0.5. When the values of the Hurst exponent lie close to 1.0, the system has long-memory dependence. The larger the H value is, the stronger the trend.

The Hurst exponent is a measure that has been widely used to evaluate the self-similarity and correlation properties of fractional Brownian noise, the time-series produced by a fractional (fractal) Gaussian process. We can describe the self-similarity process by the following equation:

$$X(at) = a^H X(t) \tag{11}$$

where a is a positive constant, and H is the self-similarity parameter, for $0 < H < 1$.

As originally defined by Mandelbrot [15], the Hurst exponent H describes (among other things) the scaling of the variance of a stochastic process $y(t)$,

$$\sigma^2 = \int_{-\infty}^{+\infty} y^2 f(y, t) dy = ct^{2H} \tag{12}$$

where c is constant.

The Hurst exponent is used to evaluate the presence or absence of long-range dependence and its degree in a time-series. For more information see [16].

The Hurst exponent (H) is defined in terms of the asymptotic behavior of the rescaled range as a function of the time span of a time series as follows

$$E\left[\frac{R(n)}{S(n)}\right] = Cn^H \text{ as } n \to \infty \tag{13}$$

where $[R(n)/S(n)]$ is the rescaled range; $E[y]$ is the expected value; n is the number of data points in a time series, C is a constant. We have used a methodology known as Rescaled Range analysis or R/S analysis. For more information see [17].

2.7 Fractal Dimension

The term "fractal" was first introduced by Mandelbrot [15]. A fractal is a complicated geometric figure that, unlike a conventional complicated figure, does not simplify it is magnified. In the way that Euclidean geometry has served as a descriptive language for classical mechanics of motion, fractal geometry is being used for the patterns produced by chaos [10].

Everybody suspects what is dimension, that is Euclidean dimension. Briefly, Euclidian dimension is given by the number of phase variables. However, for a deeper understanding of behavior of dynamical systems we must define fractal dimension. There are many specific definitions of fractal dimension. Generally, the fractal dimension, D, is a statistical quantity that gives an indication of how completely a fractal appears to fill space, as one zooms down to finer and finer scales. Let S be a set of points in a space of Euclidean dimension d. We now consider certain hypercubes of side ε, and calculate the minimum number of such cells, $N(\varepsilon)$, necessary to "cover" S. Let definite capacity dimension (Kolmogorov dimension), which is typical and common, be an example of fractal dimension:

$$D = \lim_{\varepsilon \to 0} \frac{\ln(N(\varepsilon))}{\ln \varepsilon^{-1}} \tag{14}$$

Notice that fractal dimension is a real number. A non-integer dimension does not imply chaotic dynamic, but all strange attractors must have non-integer fractal dimensions [17].

The Hurst exponent H is directly related to fractal dimension D, because maximum fractal dimension for a planar tracing is 2:

$$D + H = 2 \tag{15}$$

2.8 Correlation Dimension

In practice, capacity dimension cannot be computed easily. A different approach has been designed by Grassberger and Procaccia [18]. The method is based on the concept of correlation dimension suggested by Grassberger and Procaccia [18].

Correlation dimension (D_C) describes the dimensionality of the underlying process in relation to its geometrical reconstruction in phase space. Correlation dimension is calculated using the fundamental definition. Define the correlation integral for set of data M:

$$C(r) = \frac{1}{M(M-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{M} H(r - \|y_i - y_j\|) \tag{16}$$

where H is the Heaviside step function.

$$H(x) = \begin{cases} 0 & \text{y} < 0 \\ \frac{1}{2} & \text{y} = 0 \\ 1 & \text{y} > 0 \end{cases} \tag{17}$$

Euclidean metric is used for all calculations. When a lower limit exists, the correlation dimension is then defined as

$$D_C = \lim_{\substack{r \to 0 \\ M \to \infty}} \frac{\ln(C(r))}{\ln(r)} \tag{18}$$

3 Analysis of Electricity Price Series

3.1 Input Data

For our case study we decided to use EPEX Phelix spot index from 23.1.2012 to 13.4.2013. This is 10000 samples, which is a sufficient amount of data for our study. Descriptive statistics of the spot data are shown in the Table 1. We can see that there are extreme maximum and minimum values caused by unexpected price jumps and high standard deviation, which prove our statement about big and frequent volatility.

Table 1 Descriptive statistics of the spot prices

Mean	Median	Std dev	Skewness	Kurtosis	Min	Max	25 % qtl	75 % qtl
43,05	41,82	18,14	−2,43	41,34	−221,9	210	34,37	53,13

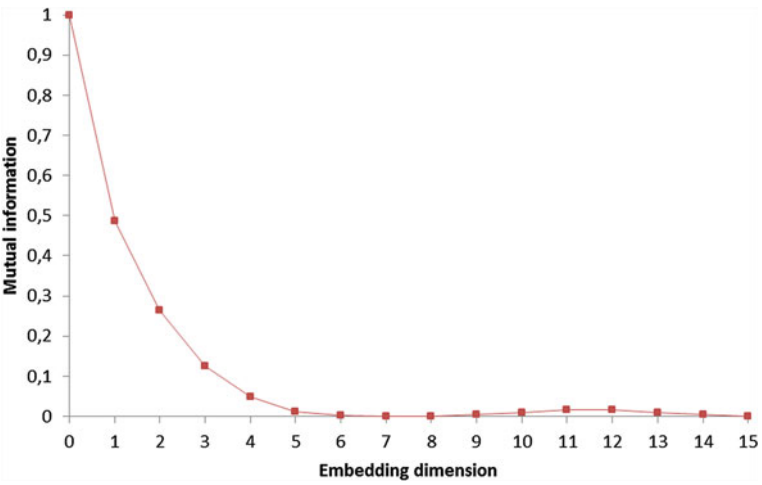


Fig. 1 Estimation of the time delay

3.2 Estimation of the Time Delay

The mutual information approach, which is described above, will be used in this chapter to determine the time delay. This variable is estimated from the graph (see Fig. 1) as the first minimum of the mutual information function. Optimal value of the time delay equals 7, which is a similar value to that of the New England Power Market time delay, which is 8. In general, there is a certain similarity between the two power markets (Central European Energy Exchange and New England Power Market [19]).

3.3 Estimation of the Embedding Dimension

The false neighbor method will be used in this chapter to determine the minimal sufficient embedding dimension. Again, this variable is estimated from the graph (see Fig. 2). The minimum embedding dimension capable of containing the reconstructed attractor is that for which the percentage of false nearest neighbors drops to zero for a given tolerance level ε. The value of the embedding dimension equals 7, which is nearly the same as the embedding dimension value of the New England Power Market, which is 9 [19].

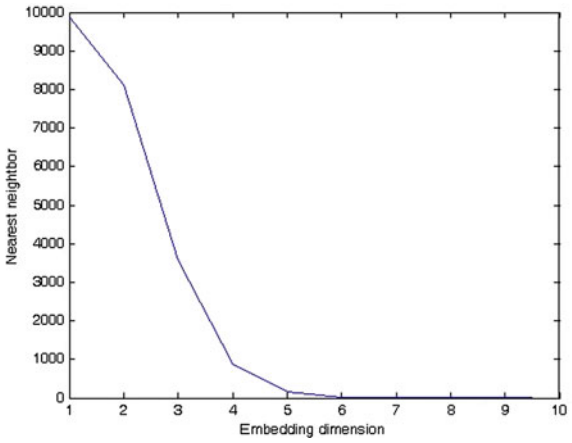


Fig. 2 Estimation of the embedding dimension

3.4 Estimation of Lypunov Exponents

The calculation of the largest Lyapunov exponent (LLE) will be made according to the theoretical background shown above, using the Rosenstein algorithm. For accurate estimation of the LLE, we need values of the embedding dimension and time delay estimated above.

A positive largest Lyapunov exponent is one of the necessary conditions for chaotic behavior. This shows that the electricity price evolution is sensitive to the initial conditions. The value of the largest Lyapunov exponent was estimated at 0.0005 for embedding dimension 7 and time delay 7. Notice that the largest Lyapunov exponent is relatively small. Consequently, the rate of the electricity price evolution is rather slow, showing that it is possible to accurately make a short-term electricity price forecast [20].

3.5 Results of the 0–1 Test for Chaos

In this chapter we calculate the correlation coefficient as was shown above. The value of the correlation coefficient was computed at 0.98. The correlation coefficient is near to 0 for non-chaotic data and near 1 for chaotic data. The value 0.98 is closer to 1. Hence we can assume chaotic behavior in the electricity price series.

3.6 Estimation of H and D_C

We have computed the Hurst exponent $H = 0.76$ according to the algorithm in Sect. 2.4. This value is in accordance with expectations. We know that the value of H is between 0 and 1, whilst real time series are usually higher than 0.5. If the exponent

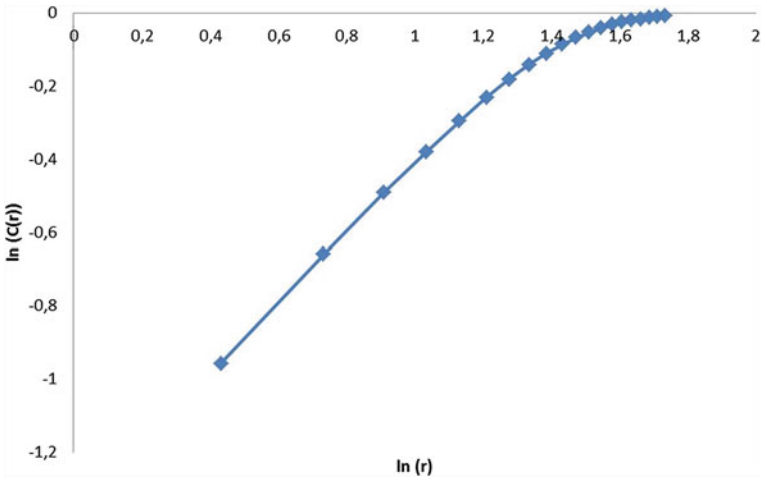


Fig. 3 The dependence of the correlation integral is on the radius in log-log graph

value is close to 0 or 1, it means that the time-series has long-range dependence. The value 0.76 is between the stochastic and deterministic process. We think, that the 0.76 value is a sufficient value for credible prediction. Now we also know the fractal dimension $2 - 0.76 = 1.24$.

The correlation dimension is calculated using the fundamental definition in Sect. 2.7. We have put the calculated data into a graph in logarithmic coordinates, and we have made a linear interpolation (cf. Fig. 3). On this basis, the correlation dimension for the small value of r can be estimated. The estimate of the correlation dimension is 1.3. As expected, the value of the correlation dimension is not an integer.

4 Conclusion

We have shown in this chapter that the electricity price series is chaotic. The estimate of the correlation dimension is 1.3 and estimate of the largest Lyapunov exponent is 0.0005. If the correlation dimension is low, the largest Lyapunov exponent is positive and the Kolmogorov entropy has a finite positive value, chaos is probably present. Then, we conducted the 0–1 test for chaos and chaos is present according to this test.

From these estimates it can be concluded that the electricity price series is chaotic. Long memory was deduced conclusively from the calculation of the values of the Hurst exponent.

References

1. Geva, B.A.: ScaleNet—multiscale neural–network architecture for time series prediction. IEEE Trans. Neural Networks **9**(5), 1471–1482 (1998)
2. Mirikitani, D., Nikolaev, N.: Nonlinear maximum likelihood estimation of electricity spot prices using recurrent neural networks. 0941–0643 Neural Comput. Appl. **20**(1), 79–89 (2011)
3. Henry, B., Lovell, N., Camacho, F.: Nonlinear dynamics time series analysis. In: Akay, M. (ed.) Nonlinear Biomedical Signal Processing, Insititue of Electrical and Electronics Engineers Inc, vol. 2, pp. 1–39. Wiley, New York (2001)
4. Takens, F.: Detecting strange attractor in turbulence. In: Rand, D.A., Young, L.S. (eds.) Lecture Notes in Mathematics vol. 898, pp. 366–381. Springer, Berlin (1981)
5. Packard, N.H., Crutchfield, J.P., Farmer, J.D., Shaw, R.S.: Geometry from a time series. Phys. Rev. Lett. **45**, 712–716 (1980)
6. Kříž, R, Chaotic Analysis of the GDP Time Series, In: Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems ISBN 978-3-319-00541-6, vol. 210, pp. 353–362 (2013)
7. Kodba, S., Perc, M., Marhl, M.: Detecting chaos from a time series. Eur. J. Phys. **26**, 205–215 (2005)
8. Fraser, A.M., Swinney, H.L.: Independent coordinates for strange attractors from mutual information. Phys. Rev. A **33**, 1134–40 (1986)
9. Kennel, M.B., Brown, R., Abarbanel, H.D.I.: Determining embedding dimension for phase space reconstruction using a geometrical construction. Phys. Rev. A. **45**, 3403–3411 (1992)
10. Cvitanović, P., Artuso, R., Mainieri, R., Tanner, G., Vattay, G., Chaos.: Classical and quantum [ChaosBook.org] Sv. version13.5. Gone with the wind press, Atlanta (2011)
11. Gotthans, T.: Advanced algorithms for the analysis of data sequences in Matlab. Master's Thesis, University of technology Brno (2010)
12. Rosenstein, M.T., Collins, J.J., Luca, C.J.: A practical method for calculating largest Lyapunov exponents from small data sets. Physica D **65**, 117–134 (1994)
13. Gottwald, G.A., Melbourne, I.: A new test for chaos in deterministic systems. Proc. Roy. Soc. A **460**, 603–611 (2004)
14. Dawes, J.H.P., Freeland, M.C.: The 0-1 test for chaos and strange nonchaotic attractors. http://www.people.bath.ac.uk/jhpd20/publications (2008)
15. Mandelbrot, B.B.: The fractal geometry of nature. W.H. Freeman and Co, New York (1983)
16. Hurst, H.E.: Long term storage capacity of reservoirs. Trans. Am. Soc. Eng. **116**, 770–799 (1951)
17. Kříž, R.: Chaos in GDP. Acta Polytech. **51**(5), 63–68 (2011)
18. Grassberg, P., Procaccia, I.: Characterization of strange attractors. Phys. Rev. Lett. **50**, 346 (1983)
19. Yang, H., Duan, X.: Chaotic characteristics of electricity price and its forecasting model. Can. Conf. Electr. Comput. Eng. (CCECE) **1**, 659–662 (2003)
20. Liu, Z., Yang, H., Lai, M.: Electricity price forecasting model based on chaos theory. In: The 7th International Power Engineering Conference, IPEC 2005. (2005)

Modeling Financial Time Series: Multifractal Cascades and Rényi Entropy

Petr Jizba and Jan Korbel

Abstract We show that a number of realistic financial time series can be well mimicked by multiplicative multifractal cascade processes. The key observation is that the multi-scale behavior in financial progressions fits well the multifractal cascade scaling paradigm. Connections with Kolmogorov's idea of multiplicative cascade of eddies in the well developed turbulence are briefly discussed. To put some flesh on a bare bones we compare volatility time series for S&P 500 stock index with a simulated multiplicative multifractal cascade processes. Qualitative agreement is surprisingly good. Salient issues, such as Codimension functions or Multifractal Diffusion analysis and its role in scaling identification are also discussed.

Keywords Multiplicative cascades · Rényi entropy · Multifractal volatility.

1 Introduction

Modeling stochastic processes in such complex multi-scale environments as financial markets is a very demanding task, that has been intensely studied since the early 20th century. Mighty impetus came in 1970 and 1980s with the advent of (multi-)fractal calculus and theory of critical phenomena. Many complex systems like strange attractors, biological and sociological systems, or financial markets are best described by their scaling behavior phrased via *scaling exponents*. Scaling and ensuing *fractal* properties of the system typically point to its inherent multi-scale complex nature. In case of multiple, or time-dependent scaling, we talk about *multi-*

P. Jizba · J. Korbel (✉)
Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Břehová 7, 115 19 Praha 1, Czech Republic
e-mail: korbeja2@fjfi.cvut.cz

P. Jizba
e-mail: petr.jizba@fjfi.cvut.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_22,

fractal systems. The formalism of the so-called *multifractal analysis* enables one to identify the presence and magnitude of various scaling exponents. The formalism of (multi-)fractal analysis, originally described in Refs. [1–3] has already many applications, for instance, in hydrology [4], fractional dynamics [5], DNA nucleotides structure [6], climate [7] and financial markets [8, 9]. It should be stressed that even within the multifractal analysis, there are many distinct approaches that are actively used for study of scaling exponents. Among these belong Multifractal Detrended Fluctuation Analysis [10], Generalized Hurst Exponents [11] and theory of Rényi entropy [12]. Particularly the Rényi entropy approach has been recently successfully applied into various financial-market contexts [13, 14]. Another very powerful concept within the multifractal analysis is a formalism of multifractal cascades [15] that enables to model realistic time series with presumed multiplicative scaling. This is quite pertinent, for instance, for understanding the volatility progression. In this chapter we focus our attention on modeling the volatility time behavior with the concept of multifractal cascades. We compare it with the more conventional methods of Multifractal Diffusion Entropy analysis. The comparison is done on the real time series of S&P 500 20-day volatility. The results obtained are discussed from both, theoretical and practical points of view, with the accent on the formation of the model series and their fit to the empirical time series. The article is organized as follows: in Sect. 2 we revise some basic aspects of multifractal analysis. Section 3 is dedicated to the multifractal cascades and their properties. In Sect. 4 we introduce the Multifractal diffusion entropy analysis and in Sect. 5 we present numerical results of volatility simulation and ensuing multifractal analysis. The final section is devoted to conclusions.

2 Multifractal Analysis

In the chapter we deal with a few approaches that enable to qualify and quantify the presence of multiscaling in financial time series. In this section we briefly remind the classical approach based on multifractal spectrum studied in Refs. [2, 3, 16]. The classical definition of multifractal spectrum of a time series is made as follows: we assume a discrete (or discretized) time series $\{x_i\}_{i=1}^{N}$ with characteristic time lag s, whose values are divided into different regions. The probability of each region p_j is defined as a ratio of points lying in the region to the total length of the series for large N, so $p_j = \lim_{N\to\infty} \frac{N_j}{N}$. The probability of each region scales with different (Lipschitz-Hölder) exponent α_j, so we have $p_j \propto s^{\alpha_j}$. It is the key assumption in the multifractal theory that α changes smoothly and thus the distribution can be assumed in the form

$$p(\alpha)\mathrm{d}\alpha = Pr[\alpha'|\alpha' \in (\alpha, \alpha + \mathrm{d}\alpha)] = \rho(\alpha)s^{-f(\alpha)}\mathrm{d}\alpha\,. \tag{1}$$

The function $f(\alpha)$, which is the scaling exponent of the distribution, is called *multifractal spectrum* and represents a fractal dimension of the subset with scaling exponent α. A dual description of multifractality is provided via the *partition function*, which is a sum of all p_i^q and its scaling exponent $\tau(q)$, the analog of free energy, reads

$$Z(q,s) \;=\; \sum_j p_j^q \propto s^{\tau(q)}\,. \tag{2}$$

The partition function is nothing else than the expected value of $(q-1)$th power of the probability distribution $\mathcal{P}$. This suggests to investigate also the generalized mean (or power mean) of $\mathcal{P}$, which is a special case of f-mean defined as $f^{-1}\left(\sum_i p_i f(p_i)\right)$, where $f(x) = x^{q-1}$. The scaling exponent D_q of the generalized mean

$$\langle \mathcal{P} \rangle_q \;=\; \sqrt[q-1]{\langle \mathcal{P}^{q-1} \rangle} \propto s^{D_q}\,, \tag{3}$$

is called a *generalized dimension* and from (2) is equal to $\frac{\tau(q)}{q-1}$. We shall note that there are some significant values of q, for which is the generalized dimension equal to other frequently known dimensions (see Ref. [12]); for $q = 0$ we get a box-counting dimension of the support, for $q \to 1$ is D_1 equal to information dimension and D_2 is equal to correlation dimension. From Eq. (3) we can express D_q as

$$D_q = \lim_{s\to 0} \frac{1}{q-1} \frac{\ln Z(q,s)}{\ln s} = \lim_{s\to 0} \frac{S_q(s)}{\ln s}\,, \tag{4}$$

where S_q is a Rényi entropy associated with the distribution $\mathcal{P}$, which is one-parametric generalization of Shannon entropy (for $q \to 1$). With different choice of q, regions with high or low probability are accentuated.[1] The relation between $f(\alpha)$ and $\tau(q)$ can be easily obtained from the fact that the partition function is equal to

$$Z(q,s) \;=\; \int \mathrm{d}\alpha\, \rho(\alpha) s^{-f(\alpha)} s^{q\alpha}\,, \tag{5}$$

and from the stationary phase approximation we conclude that the relation is

$$\tau(q) = q\alpha(q) - f(\alpha(q))\,, \tag{6}$$

where $\alpha(q)$ is the value that extremizes $q\alpha - f(\alpha)$. When we also consider a differentiability of scaling exponents, then $\alpha = \frac{\mathrm{d}\tau(q)}{\mathrm{d}q}$, so Eq. (6) represents nothing else than the Legendre transform of $f(\alpha)$.

[1] We shall note that for $q < 0$ the definition of entropy can be problematic (e.g. can exhibit instabilities, see e.g. Ref. [12]), and therefore shall deal only with $q \geq 0$.

3 Multiplicative Cascades

The theory of multiplicative processes was firstly formulated by A.N. Kolmogorov in 1940's seminal paper [17], in connection with the theory of the well developed turbulence. The model of whirls assumes that larger eddies are composed of eddies on smaller scale with some typical distributions. The idea behind multiplicative cascades is that the typical quantity (i.e. the turbulence dissipation rate in hydrodynamics or volatility in finance) measured on some scale is compound of the some quantity measured on the smaller scale in subregions of the original region. For this end, we define a sequence of typical scales $r_0 > r_1 > \cdots > r_n$ and typical multipliers $l_n < 1$, which are the ratios of two successive scales, so $r_n = r_0 \prod_{j=1}^{n} l_j$. Let us denote a characteristic quantity of the system as E, and let $\epsilon(x)$ be a density of this quantity, so the total value over the region Ω is equal to

$$E_\Omega = \int_{x \in \Omega} \epsilon(x) \mathrm{d}x \ . \tag{7}$$

When the quantity is modeled by a multiplicative cascade, we define it on scales r_n (similarly as the scale itself) as a product of multipliers M_j, so

$$E_{r_n} = E_{r_0} \prod_{j=1}^{n} M_j \ , \tag{8}$$

which for $n \to \infty$ defines the density $\epsilon(x)$. Thus, the cascade is determined by the multipliers l_j and M_j. The multipliers M_j (and possibly also l_j) can be also assumed as random variables. When the variables M_j are independent of j, i.e. the multiplier remains the same for every scale, we can observe self-similar behavior and the cascade is fractal. We shall note that if we need to distinguish a concrete realization of M_j on the region i, we denote it as $M_{j,i}$.

There are many models of multiplicative cascades, we mention only the most popular. In the original work, Kolmogorov considered the distribution to be isotropic [17], so the only quantity that is relevant for the model is $\langle E_{r_0} \rangle$. The other popular model is a log-normal model, where the multiplier has log-normal distribution, or the β-model, where a fraction of M_j. is nonzero and equal, while the rest is equal to zero. These models usually describe complex systems insufficiently, so therefore we turn our attention to a different class of cascades which produces non-trivial multiscaling properties.

3.1 Multifratal Cascades

Multifractal cascades is a class of cascades with a multiple scaling exponents. The complexity of the cascade enables to model real systems, as chaotic systems [15], rainfalls [18] or prices on financial markets [8]. There are several models of

multifractal cascades. The simplest model, *binomal cascade*, is a deterministic version of multifractal cascade and is a springboard for further generalizations. In every step is the original region divided into two different regions ($l_j = \frac{1}{2}$), and M_{n-1} is divided into two subregions, so $E_{n,1} = pE_{n-1}$, $E_{n,2} = (1-p)E_{n-1}$, so $M_1 = p$ and $M_2 = 1-p$. When we assume a division into more than two regions ($l_j = \frac{1}{n}$), we become a *multinomial cascade*. The important property of this cascade is that it is *conservative*, i.e. $\sum_i M_{n,i} = 1$. The *microcanonical cascade* is a straightforward generalization of the binomial cascade, where allow to interchange M_1 and M_2, so $M_1 = p$ with probability $\frac{1}{2}$ or $M_1 = 1-p$ with the same probability. Unfortunately, in this case are M_1 and M_2 not independent. The independency of M_1 and M_2 is possible to reach by weakening the condition of conservation, i.e. by replacing it by the *statistical conservation*, which means that only the expectation value is conserved, i.e.

$$\langle \sum_i M_{n,i} \rangle \;=\; 1\,. \tag{9}$$

Together with the independence of random variables we obtain $\langle M_j \rangle = l_j$.

3.2 Codimension Function

An alternative description of multifractality can be provided via singularity spectrum and codimension function (cf., e.g., Ref. [15]). Let us assume that the typical scale is defined through ratio λ, so $r_n = \frac{r_0}{\lambda^n}$. Let us also have a multiplier $\mathcal{M}$, such that the some typical quantity E is on every scale iteratively compound of itself, measured on smaller scales, hence

$$E_{r_n} \;=\; E_{r_{n-1}} \cdot \mathcal{M}\,. \tag{10}$$

We suppose that the moments of $\mathcal{M}$ scale as $\langle \mathcal{M}^q \rangle \sim \lambda^{K(q)}$ and from that we have

$$\langle E_{r_n}^q \rangle \;\sim\; \Lambda^{K(q)}\,, \tag{11}$$

where $\Lambda = \lambda^n$. It is further natural to extend these special values of Λ to any positive scale. On the other hand, $E_{1/\Lambda}$ can be described via its *singularity spectrum*. We consider the probability distribution that $\mathcal{M} \sim \lambda^{\gamma'}$, where $\gamma' > \gamma$, in a form

$$Pr[\gamma' \geq \gamma] \sim \rho(\gamma)\lambda^{-c(\gamma)}\,, \tag{12}$$

where $c(\gamma)$ is the singularity spectrum. The relation between $K(q)$ and $c(\gamma)$ is (similarly for $f(\alpha)$ and $\tau(q)$) easily obtained from the expression of $\langle \mathcal{M}^q \rangle$, i.e.

$$\langle \mathcal{M}^q \rangle \;=\; \int \mathrm{d}\lambda\; \lambda^{q\gamma}\tilde{\rho}(\gamma)\lambda^{-c(\gamma)}\,. \tag{13}$$

From the stationary phase approximation we finally obtain that

$$K(q) \;=\; q\gamma(q) - c(\gamma(q))\,. \tag{14}$$

The relation to the classical multifractal analysis is determined by the relation between the theoretical distribution γ' and empirical distribution obtained from reproducing copies of given cascade, so

$$Pr[\gamma' > \gamma] \;=\; \lim_{N_\lambda \to \infty} \frac{N_\lambda(\gamma' > \gamma)}{N_\lambda}\,, \tag{15}$$

where $N(\gamma' > \gamma)$ is a number structures with scaling exponent larger than γ and N_λ is total number of regions. It is easy to show that $N(\gamma' > \gamma)$ scales as $\lambda^{f(\alpha(\gamma))}$ (see again, e.g., Ref. [15]), so from that we finally obtain relations:

$$f(\alpha(\gamma)) = D - c(\gamma)\,, \tag{16}$$

$$\alpha = D - \gamma\,, \tag{17}$$

$$\tau(q) = D(q-1) - K(q)\,. \tag{18}$$

The usage of codimension formalism can be advantageous, because "codimension quantities" do not depend on the dimension, so especially in higher-dimensional problems we observe similar spectrum irrespective of the dimension involved. We shall also note, that in case of monofractal series we obtain the well known relation between fractal dimension and Hurst exponent, which plays here the role of codimension function in two-dimensional $x - t$ space.

4 Multifractal Diffusion Entropy Analysis (MFDEA)

In order to measure fractal dimension and multifractality, there have been developed many methods including, e.g., Detrended Fluctuation analysis [6, 10], Generalized Hurst exponent [11] or Multifractal Wavelet analysis [19]. Here we briefly describe yet another method for multifractality analysis, namely Multifractal Diffusion Entropy analysis (cf. also Ref. [13]). The latter is reasonable particularly in situations, when it is reasonable to expect that system produces heavy-tailed distributions. Let us begin with a monofractal version. In this case, the method is based on a functional ansatz for the self-similar form of the probability density function

$$p(x,t) \;=\; \frac{1}{t^\delta} F\left(\frac{x}{t^\delta}\right). \tag{19}$$

The scaling exponent δ can be read off from Shannon entropy

$$S(t) \;=\; -\int \mathrm{d}x\; p(x,t)\ln[p(x,t)]\,, \tag{20}$$

as $S(t) = A + \delta \ln t$. If we generalize Shannon entropy into the one-parametric class of Rényi entropies, we obtain the whole class of scaling exponents parametrized by q, namely

$$S_q(\tau) \;=\; B_q + h(q)\ln\tau\,. \tag{21}$$

The method is based on probability density estimation by *Fluctuation collection*, which means that from noise-like series $\{\xi_i\}_{i=1}^{N}$ we create fluctuation collections $x_\tau(t) = \sum_{i=1}^{s} \xi_{i+t}$. After this all values of $x_\tau(t)$ are divided into boxes of length ϵ and the probability distribution is estimated as a relative frequency of each box, so $p_i(\tau) = \frac{N_i(s)}{N-\tau+1}$, where $N_i(s)$ is number of $x_\tau(t)$ that fall into the ith box. Finally, the estimated entropy has the form

$$\hat{S}_q(\tau) = \frac{1}{q-1}\ln\sum_i [p_i(\tau)]^q\,. \tag{22}$$

5 Numerical Simulation of Volatility as a Multifractal Cascade

An important quantifier of riskiness inherent in a given financial time series is undoubtedly *volatility*. Modeling volatility is typically the key element in any risk-rating analysis of financial-time progressions. The approach firstly described by Mandelbrot, Calvet and Fisher (see, e.g., Refs. [8, 20]) models a volatility series as a multifractal cascade with some given (usually discrete) distribution. Here we simulate this multifractal cascade and compare with an empirical time series, namely 20-day volatility series of the financial index S&P 500. For the simulation we have chosen a diadic partition and a multiplier $\mathcal{M}$ with three values and probabilities drawn from the distribution $Binomial(2, p)$. The values were normalized, so that $\langle\mathcal{M}\rangle = 1/2$, yielding the so-called conservative multiplier. The simulated series and volatility of S&P 500 is displayed on Fig. 1. Both mentioned series were normalized, so that the scale on the picture is in multiples of the mean of the series. The standard deviation of the renormalized S&P 500 volatility series is 0.65, whilst the standard deviation of the simulated series is 0.72. It should, however, be born in mind, that the mean and standard deviation can shed a very little light on the complexity of the series, which can comprise presence of seasonality, sudden jumps and spikes or nontrivial volatility behavior (e.g., "volatility smile"). These different "regimes", resulting from the nature of the system can be successfully investigated by analysing the spectrum of ensuing scaling exponents, each representing one such mode or time scale. We shall remark that the simulated time series can be used to create the series of daily returns, when we consider that for daily returns holds $r_t = \sigma_t \eta_t$ where σ_t is simulated volatility and η_t is a white noise.

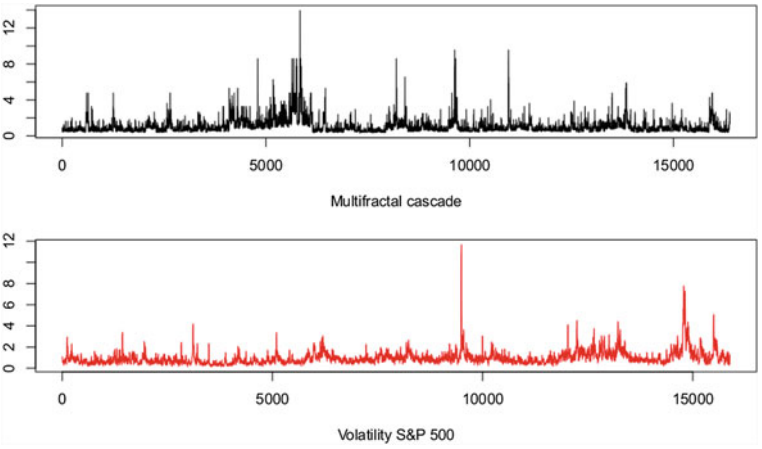


Fig. 1 Simulated volatility modeled as a multifractal cascade and 20-day volatility of S&P 500 index. Both series contain approximately 16000 entries

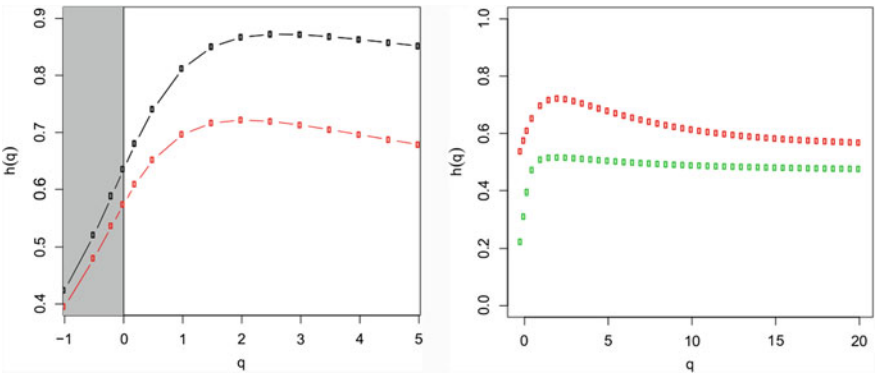


Fig. 2 *Left* scaling exponent $h(q)$ of both series. The region of $q < 0$ is shaded, because the entropy has there instabilities. *Right* scaling exponent h(q) of 20-day volatility of S&P 500 (*red*) and daily returns of S&P 500 (*green*)

In the second part of the presented analysis we estimate scaling exponents $h(q)$ of Rényi entropy. This is done with the aim to measure the multifractality and compare it for both series. As already mentioned, the scaling exponents provide an important signatures of the intricate multi-scale behavior in time series and thus can shed a non-trivial light on time-scale behavior in realistic financial markets. Figure 2 shows the exponent $h(q)$ for both series. We observe that the multiplicative cascade has wider spectrum of exponents, with highest values close to 0.9, whereas the maximal exponent of S&P 500 volatility series is just around 0.7. That means that for more precise simulation the distribution of the multiplier $\mathcal{M}$ should be adjusted in order to get a spectrum that would be more similar to the real one. Nevertheless, spectra

of different time series observed in the financial markets can be shifted up or down, and sometimes quite remarkably (as in Ref. [13]). The other interesting thing is the behavior of the exponent for large q's. On graph of Fig. 2 we compare the spectrum of volatility and daily returns of S&P 500. While for returns we observe a fast convergence to a single value of scaling exponent, in case of volatility we observe a rapid increase followed by a moderate decrease and slow convergence. At the moment it is not clear whether this rapid behavior points to yet unknown complexity in volatility time-series which goes beyond a simple multifractal paradigm or whether it is a mere numerical artefact. If the first explanation would turn out to be the case, it could open new vistas for mathematical modeling of multiple scale behavior and successive deeper understanding of inherent complexity in financial time series.

6 Conclusions

Modeling complex multi-scale time series, such as volatility time series in financial markets, requires novel mathematical techniques that are able to appropriately accentuate typical scaling characteristics observed in realistic financial markets. In this chapter we have put forward idea that an approach based on multiplicative multifractal cascades is capable of creating time series that correctly mimic scaling behavior seen in a number of empirical financial time series. The choice of an appropriate cascade multiplier is crucial for the method in question, and as we have illustrated in this chapter, methods of multifractal analysis (as, e.g., Multifractal Detrended Fluctuation analysis or Multifractal Diffusion Entropy analysis) are particularly useful in this context. Apart from use in quantitative finance, the method proposed reinforces also the links between the multiplicative multifractal cascades and the multi-scale behavior in a wide class of complex dynamical systems. Further more quantitative study supporting the usefulness of multifractal analysis in time series in finance is left for a future publication. Finally, we hope that the article will stimulate the reader from other fields to look more closely into this fascinating subject.

Acknowledgments Authors want to thank for the financial support provided by the Grant Agency of the Czech Republic, grant No. GCP402/12/J077, and the Grant Agency of the CTU in Prague, grant No. SGS13/217/OHK4/3T/14.

References

1. Mandelbrot, B.: Self-affine fractals and fractal dimension. Physica Scripta **32**, 257–260 (1985)
2. Hentschel, H., Procaccia, I.: The infinite number of generalized dimensions of fractals and strange attractors. Physica D **8**, 435–444 (1983)
3. Harte, D.: Multifractals: theory and application. Chapmann & Hall/CRC, Boca Raton (2001)
4. Koscielny-Bunde, E., Kantelhardt, J., Braun, P., Bunde, A., Havlin, S.: Long-term persistence and multifractality of river runoff records: detrended fluctuation studies. J. Hydrol. **322**, 120–137 (2006)

5. Klafter, J., Lim, S.L., Metzler, R.: Fractional Dynamics. Imperial College Press, London (2011)
6. Peng, C.K., Buldyrev, S., Havlin, S., Simons, M., Stanley, H., Goldbergerz, A.: Mosaic organization of DNA nucleotides. Phys. Rev. E **49**, 1685–1689 (1994)
7. Lovejoy, S., Schertzer, D.: The Weather and Climate: Emergent Laws and Multifractal Cascades. Cambridge University Press, Cambridge (2013)
8. Mandelbrot, B., Calvet, L., Fisher, A.: A multifractal model of asset returns. (Cowles Foundation Discussion Papers)
9. Mandelbrot, B.: Multifractals and 1/f Noise: Wild Self-Affinity in Physics (1963–1976). Springer, New York (1999)
10. Kantelhardt, W., Zschiegner, S., Koscielny-Bunde, E., Havlin, S., Bunde, A., Stanley, H.: Multifractal detrended fluctuation analysis of nonstationary time series. Physica A **316**, 87–114 (2002)
11. Morales, R., Di Matteo, T., Gramatica, R., Aste, T.: Dynamical generalized hurst exponent as a tool to monitor unstable periods in financial time series. Physica A: Stat. Mech. Appl. **391**, 3180–3189 (2012)
12. Jizba, P., Arimitsu, T.: The world according to Rényi: thermodynamics of multifractal systems. Ann. Phys. **312**, 17–59 (2004)
13. Huang, J., et al.: Multifractal diffusion entropy analysis on stock volatility in financial markets. Physica A **391**, 5739–5745 (2012)
14. Jizba, P., Kleinert, H., Shefaat, M.: Rényi information transfer between financial time series. Physica A **391**, 2971–2989 (2012)
15. Schertzer, D., Lovejoy, S., Schmitt, F., Chigirinskaya, Y., Marsan, D.: Multifractal cascade dynamics and turbulent intermittency. Fractals **5**, 427–471 (1997)
16. Mandelbrot, B.: Heavy tails in finance for independent or multifractal price increments. In Rachev, S.T. (ed.) Handbook of Heavy Tailed Distributions in Finance. Elsevier, Amsterdam (2003)
17. Kolmogorov, A.N.: The local structure of turbulence in incompressible viscous fluids at very large reynolds numbers. Dokl. Akad. Nauk SSSR. **32**, 16–18 (1941)
18. Tessier, Y., Lovejoy, S., Schertzer, D.: Universal multifractals: theory and observations for rain and clouds. J. Appl. Meteorol. **32**, 223–223 (1993)
19. Muzy, J.F., Bacry, E., Arneodo, A.: Multifractal formalism for fractal signals: the structure-function approach versus the wavelet-transform modulus-maxima method. Phys. Rev. E **47**, 875 (1993)
20. Calvet, L., Fisher, A.: Multifractal Volatility: Theory Forecasting and Pricing. Academic Press, Amsterdam (2008)

The Global Multi Factor Model of Seismic Activity: Priorities

Natalia P. Bulatova

Abstract It is shown that catastrophic processes of the Earth occur at simultaneous action of several groups of factors that include external global space influences (the Sun and the Moon) and internal geological influences, which provide the condition in strong earthquake area. The chaotic space-time distribution of earthquakes tells us about this. On the basis of the analysis of events occurring during 110 years it is possible to assert that on an average quantity the number of earthquakes with magnitude $M \geq 7$ has not increased. From the other hand, time intervals when the time distribution of earthquakes has periodic character are found. The purpose of this work is to reveal time intervals, during of which there is predominant influence of the separate factor. In order to do this we use statistical methods. The problem is complex because the ensemble of factors is impacting. By assuming a simultaneous influence of several factors (the Sun, the Moon and geological conditions) we performed the statistical division of the time series consisting of >500 events (during 1973–2005) into groups, in which one can see the influence of separate factors. It has been shown that the geological factor and the cyclic character of the Moon influence play the main role in appearance of the seismic activity.

Keywords Spatial-Temporal modeling · Earthquake · Solar cycle · Lunar cycles · Trends

1 Introduction

It is a long time since the interest of scientists is chained to research the nature of occurrence of earthquakes, which as a rule take place suddenly and lead to a great disasters: to destruction of buildings, death of people etc. [1–3]. If we could predict

N. P. Bulatova (✉)
Schmidt Institute of Physics of the Earth of Russian Academy of Sciences, Moscow, Russia
e-mail: n.p.bulatova@mail.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_23,

exact places and time of global natural accidents (strong earthquakes, a tsunami etc.) we would have a possibility to be prepared for these events and to avoid victims.

After many years of failures in forecasting the reason becomes clear: seismic processes develop as a result of joint action of two principal types of factors: endogenous, in the Earth (geological) and external (space). For the successful forecast it is necessary to consider seismic processes as multifactor and to be able to reveal the contribution of each of them. Without the solution of a problem of a continuous monitoring by a wide network of seismic stations it is impossible to give reliable forecasts of accidents and earthquakes. Monitoring of seismic modes of extensive territories costs too much. It is the reason why catalogues and databases are not complete and have “admissions" in time and space. Interdisciplinary approach of studying could help to discover spatial-temporal path of the solution of this problem.

For this purpose the author proposed a research tool [4–6], the algorithm which has been further developed as the 3D spatio-temporal technology [7]. Its use is allowed to combine modern astronomical and geological data to the one system. In addition it was necessary to prioritize in considering the impact of the possible factors. Some preliminary results of this analysis of time series of seismic events $M \geq 7$ connection with the lunar and the solar cyclical effects were presented by the author in 2004 [4].

Scientists Shirokov and Serafimova [8–10] have reported about interrelation of the 19-year lunar tides and 22-year Hail’s solar cycles with strong earthquakes and thus about long-term seismic forecast. For the first time in their paper they are proposed the forecast of strong volcanic eruptions of Pacific belt for the next 20 years. The forecast was made for the eight regions Pacific belt on the basis of the method of phase pathways that has been developed by authors [9, 10]. A distribution of strong eruptions in Pacific belt for data of observations in XVIII–XIX centuries has been studied for the two-dimensional phase plane. Coordinates of this plane were both in the phase of 19-years lunar tide and the phase of 22-years magnetic solar cycle. The activity is presented as sub-meridian section and it is highlighted on the border of the continental plates in the coastal zones of the Pacific Ocean. Seismic and volcanic activity is continued in some areas and it is constantly connected mainly with passing of geological and geochemical processes in the Earth’s interior. Therefore the method of phase pathways uses individual algorithms to investigate such areas.

Many another authors report that astronomical effects can have an influence, but the probability of an earthquake is built by many factors [11, 12], therefore they can discussed the information in broad terms only. The articles about the influence of electromagnetic fields of the Sun are interesting for discussion [12–14].

2 Multifactor Model. Global Approach to Spatial-Temporal Study of the Seismic Activity

It is clear that the probability of an earthquake is made up of many factors, most of them are geological, but astronomical effects can have an influence too. We cannot express the probability of a distribution of global seismicity in the form of mathematical calculations, but we can try to test the hypothesis that geological and celestial

bodies have an influence on it. By giving priorities to the multi-factor model we do the structuring of data. We analyze the results of several tests by using the world-known archives and international database [15–17]: (1) seismic data, (2) astrometric data, (3) solar activity.

2.1 The Used Data and Parameters

All data were distributed in time series with the time step-interval that is equal to one year.

Seismic data._Seismic events have occurred with a magnitude ranging from 7.0 to 10.0 ($M \geq 7$) during the next time periods (a) 1900–2010, (b) 1973–2007 years. We used data from the U.S. Geological Survey's Earthquake Data Base directory that presents the distribution of earthquake epicenters in the world [15]. In this research, the data set that consists of more than 500 seismic events have been used

Astrometric data._Lunar astrometric data for the time periods (a) 1900–2010, (b) 1973–2010 years have been analyzed. In this study, more than 16500 events from 1973 to 2002, including a full 18.6 year cycle of lunar motion between 1982 and-2001 were investigated.

For the time period 1982–2006 year. the author analyzed time rows of variations in the astronomical date δ(T) that describe declinations of the Moon [16]. Variations δ(T) have been investigated step-by-step in order to understand the dynamics of relative motion in the Sun–Earth–Moon system as changes in the angular distances bode with respect to the Earth's equatorial plane (Fig. 1a). The 12-hourly time- rows of astronomical date, which were collected by the author over the 20-year period (1982–2002) were visualized as plots of declination δ against time T.

Graphic presentation of the time rows of variations in the declinations δ_1 of the Sun and δ_2 of the Moon over 1982–2006 year. (Fig. 1b). Direct lines correspond to the values of δ_k for the singular points of the both plots corresponds to maximums, minimums, etc. of these plots: $\delta_{NA} = 28,5°$; $\delta_{NT} = 23,5°$; $\delta_{NB} = 18,5°$; $\delta_0 = 0,0°$; $\delta_{SB} = -18,5°$; $\delta_{ST} = -23,5°$; $\delta_{SA} = -28,5°$.

The cyclic variation of the lunar declination δ_c(t) is presented by the envelopes LS and LN of the region of periodic variations in the astronomical parameter of the Moon in the time plot. The envelopes NT, ST for the variations lunar δ intersect the LS, LN and divide the region of variations in the lunar declination into two parts: part A (1982–1992 year.) and part B (1993–2002 year.). The rates $\Delta\,\delta\,\Delta$T of changes in δ these parts differ substantially. For the period under consideration (1970–2007), which includes an 18.6 year cycle of the Moon's relative motion to the Earth, the author analyzed time series periodic variations (lunar period ~27 day) in the astronomical data δ(T): declinations of the Moon [18].

In order to obtain cyclic trend of the Moon (18.6 year.) cyclic movement we have found envelops around LS and LN (Fig. 2) by carrying out of two time series samplings of angular distance δ(T) fluctuation extreme (maximum and minimum) values with respect to the Earth equator (in the Northern and the Southern hemispheres respectively).

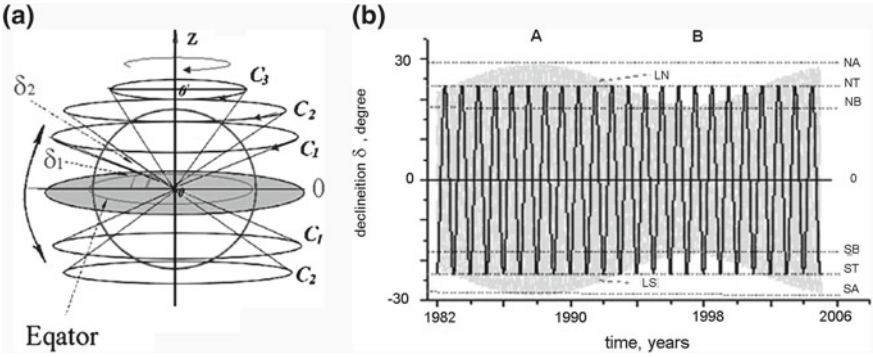


Fig. 1 **a** The three-dimensional dynamic model that was developed by the author [4]. The equatorial plane is colored in a *gray*. The *circle* denotes the cross-section of the Earth's body. C_1C_2 and C_3 denote the motion of the Sun, the Moon and other celestial body respectively, **b** Graphic presentation of the time row of variations in the declinations the δ of the Moon (field corresponds to 16500 data) and of the Sun (*black* quasi-sinusoidal curve contains 8300 data) over 1982–2006 year

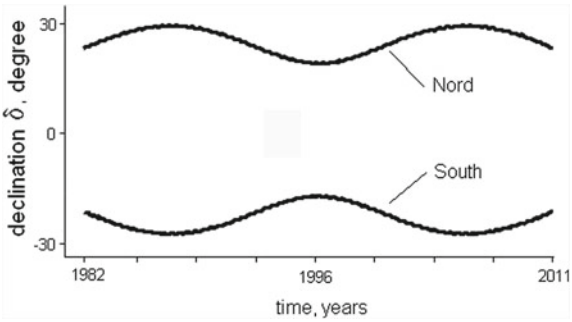


Fig. 2 The discrete values obtained at sampling have been united in curves and two quasi-sinusoid Curves are formed as the sum of two harmonics: one, the main, (with a period of ≈18.6 years), and another-additional (with a period of ≈0.5 years)

In order to form of the cyclic trend we used the approach is similar to those that described in [19]. We choose an interval in length of 9.3 years (approximately one half of the dominating harmonic period) and carry out calculations for the curve of the Moon cyclic movement with respect to the Earth. The additional harmonic disappeared.

The solar activity data: yearly average distribution 11- and 22-years Wolf's and Haile's cycles sunspots number between 1900 and 2010 [17] (Fig. 3).

3 Modeling. Research Method

We have estimated the parameters of a linear trend for the whole interval between 1970–2007 years and have subtracted it from the analyzed temporal row of strong earthquakes—N (see Fig. 4), (Fig. 5), (Fig. 6).

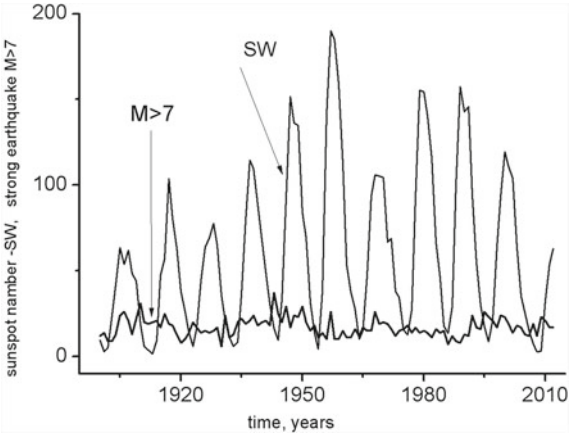


Fig. 3 The data about seismic M>7 and solar activity between 1900 and 2010 years

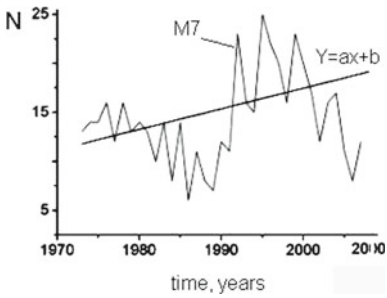


Fig. 4 The time series of strong earthquakes M $\geq$ 7 ($N_{M\geq 7} = 524$) that has been analyzed for period 1973–2005 years. The linear trend N_r for the entire time interval is shown by the direct line

The parameters of the linear trend N_r for the entire time interval were evaluated, as

$$y = ax + b, \tag{1}$$

where $a = 0.210$; $b = 11{,}159$, correlation coefficient $Q_{Nr} = 0.9071$ (Fig. 4)

$$\overline{N} = N_{M\geq 7} - N_r \tag{2}$$

The comparison procedure for $\overline{N}$ was performed with lunar cyclic trend N_c, obtained from the curve $\delta(T)$ on the basis of the linear trend (Fig. 5):

$$Y_c = b_c x + a_c \tag{3}$$

where $b_c = 0.350$; $a_c = 19{,}487$.

$$N_s = \overline{N} - N_c \tag{4}$$

The residue N_s was obtained by subtraction (4). The compare it with the 11-year cyclic changes parameter of solar activity SW (cyclical trend) showed (Fig. 6): correlation coefficient is $Q_{Ns} = (-0.32) = -0.6013$.

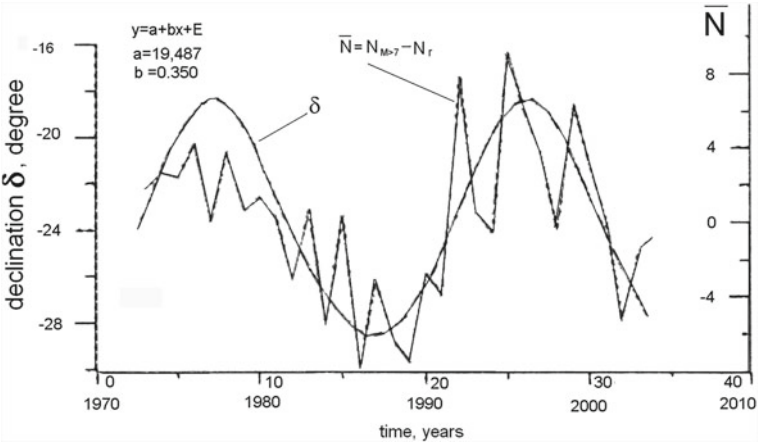


Fig. 5 A comparison of the rest of the analyzed time series of strong earthquakes of M ≥ 7 after subtraction of the linear trend *Nr* and a lunar cyclic trend for the period of 1973–2005 for the whole time interval

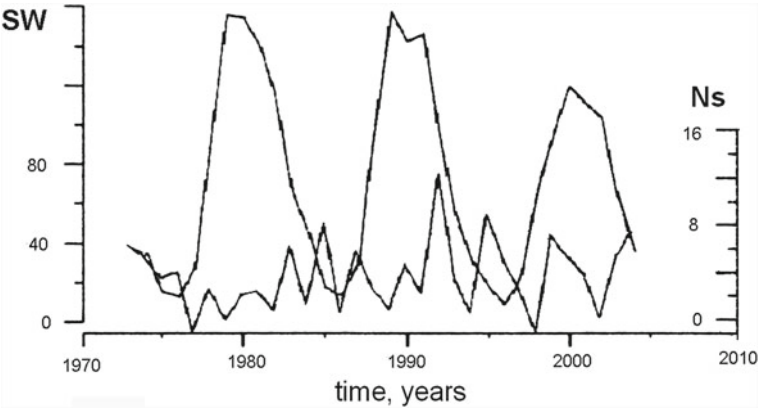


Fig. 6 A comparison of the rest of the analyzed time series N_s of strong earthquakes M ≥ 7 after subtracted cyclical trend N_c and Wolf's average annual for period 1973–2005 years for the entire time interval

As a result of this investigation the time series of strong earthquakes N was represented:

$$N_{M\geq 7} = N_r + N_{lc} + N_s, \tag{5}$$

as sum of three trends temporary earthquake rows, in accordance with which the estimates of the influence of the corresponding factor was given.

4 Conclusion

After each catastrophic event scientists analyze their reasons. However the role of many factors that simultaneously influence on the preparation and the start of earthquakes is often veiled. It is necessary to search solution in well-taken priorities of factors in an interdisciplinary area of science. In this work the author investigated the time rows of strong earthquakes and the following priorities have been chosen: a geological condition of an environment, tidal action of a gravitational field of the Moon (cyclic 18.6 years) and change of the solar activity (11 and 22 solar cycles).

By means of consecutive subtraction of linear and cyclic trends from time series(524 earthquake in the time interval 1973–2010 year) one can see that the row can be represented as three temporary earthquake rows, which have different originating sources, namely: an internal- the geological one inside the Earth (the linear trend) and the external one—the cosmic (the cyclic). Analyzing earthquake rows we conclude that the cyclic Moon impact can be clearly selected.

Let us note that for a power class of the earthquakes $M \geq 7$ the absence of correlation between the seismic and the solar activity has been obtained. Maximums of the earthquake temporary rows (in years) correspond to minima of phases of the solar activity. The similar conclusion can be also made for the earthquakes of the power classes of $M\ 5,0 \div 5,9$ and $M\ 6,0 \div 6,9$.

Acknowledgments The author expresses his gratitude to Prof. Dr. A.A. Lyubushin and Dr. V.L. Bychkov for help and consultations.

References

1. Simpson, J.F.: Solar activity as a triggering mechanism for earthquakes. Earth Planet. Sci. Lett. **3**(5), 417–425 (1968)
2. Sytinsky, A.D.: On influence solar activity to seismicity of the Earth. DAN SSSR. **208**, 1078–1081 (1973)
3. Sytinsky, A.D.: On earthquakes and solar activity bond. Phys. Earth. **3**, 110–112 (1991)
4. Bulatova, N.P.: Spatial-temporal analysis of the Earth's seismicity (in Russian). Ph.D. thesis, 150. Moscow (2004)
5. Bulatova, N.P.: On the problem of solar neutrino tomography of the Earth: scanning geometry. Izv. Phys. Solid Earth **35**(2), 150–160 (1999)
6. Bulatova, N.P.: The latitudinal distribution of terrestrial seismicity in relation to the locations of the Sun and Moon. Volcanol. Seismol **2**, 57–78 (2005)
7. Bulatova, N.P.: Three-dimensional spatio-temporal modeling of geophysical events and the movement of celestial bodies. JCS. **20**(3), 215–227 (2012)
8. Shirokov, V.A., Serafimova, Yu.K.: On the relationship of 19-year lunar tides and 22-year solar cycles to strong earthquakes and a long-term seismic forecast for regions of the northwest part of the pacific belt. Herald of the kraesc. Earth sci. **8**(2), 120–133 (2006). http://www.kscnet.ru/kraesc/index.html
9. Shirokov, V.A., Serafimova, Yu.K.: The Forecast of strong volcanic eruptions of pacific belt for the next 20 years on the basis of applying of a method of phase pathways. Herald of the kraesc. Earth sci. **2**(12), 154–163 (2008). http://www.kscnet.ru/kraesc/index.html

10. Serafimova, Yu.K., Shirokov, V.A.: The forecast of strong volcanic eruptions, volcanic eruptions and tsunami wave on the basis study both the of 18.6-year's lunar tide and of 22-year's magnetic solar cycle. Seismological and Geophysical Research on Kamchatka, pp. 305–328. New book, Petropavlovsk-Kamchatsky (2012)
11. Nesterenko, P.P., Stovas, M.V.: Change the gravitational field as one of the causes of seismicity of the earth. In: Geophysics and Astronomy, pp. 85–92. Naukova Dumka, Kiev (1963)
12. Belov, S.V., Shestopalov, I.P., Harin E.P.: On the relationship of endogenous activity of land with solar and geomagnetic activity. Rep. Acad. Sci. **428**(1), 1–4 (2009)
13. Altgauzen, N.M.: On the correlation between geomagnetic disturbances and seismic activity of the earth. Geomag. Aeron. **14**(4), 698–701 (1974)
14. Sobolev, G.A., Shestopalov, I.P., Harin, E.P.: Geoeffective solar flares and seismic activity of the Earth. Phys. Earth **7**, 85–89 (1998)
15. Global Hypocenters Data Base, National Earthquake Information Center, U.S. Geological Survey, Denver, 1973 to present
16. The Yearbook of Astronomical Observations, St. Petersburg, Russia: Institute of Applied Astronomy, Russian Academy of Sciences (1982–2010)
17. http://ngdc.noaa.gov/. Catalog FTP/STP/SOLAR_DATA
18. Podobed, V.V., Nesterov, V.V.: General Astrometry, pp. 552. Nauka, Moscow (1975)
19. Lyubushin, A.A., Klyashtorin, L.B.: Short term global Dt prediction using (60–70)-years periodicity. Energy and Environment, vol. 23, pp. 1. Multi-science publishing co. ltd., Essex, UK (2012)

Modeling Spatio-Temporal Dynamics of Taiga Boreal Forest

Andrey Lyushnin and Dmitry Bratsun

Abstract The simple three variable evolutionary model of boreal forest of Perm region has been proposed. The model is built as a complex system, where each population is represented by individual trees competing for solar light. Other factors taking into account are growth rate, seed dispersal and mortality. The parameter values used in the model were calibrated from the information available for Perm forests. This work has a fundamental aspect because a formation of dynamical macroscopic patterns in ecological systems attracts great interest of researchers. In addition, the proposed model can have many applications for more effective forest management.

Keywords Forest evolution · Cellular automata · Individual-based models · Spatio-temporal behaviour

1 Introduction

As it is known, the spontaneous formation of spatial and temporal structures is the general law of functioning of complex ecosystems. Study of such structures is one of the central problems of the modern ecology [1–3]. In plant communities a formation of spatially ordered structures which are inhomogeneous in concentration and qualitative composition of the biomass is also observed [3–5]. It should be noted that these structures are cooperative in its nature and arise due to the interaction at the level of individual plants. They should be distinguished from the structures which originate due to the morphological reason or natural landscape factors [5]. As an example of such self-organization is a "tiger bush", which is a patterned

A. Lyushnin (✉) · D. Bratsun
Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia
e-mail: andry@pspu.ru

D. Bratsun
e-mail: dmitribratsun@rambler.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_24,

vegetation community consisting of alternating bands of trees or shrubs, separated by bare ground or low herb cover, that run roughly parallel to contour lines of equal elevation [6]. This phenomenon can be met in the arid areas of Africa, Australia and North America. In recent years a number of studies have shown that the regular pattern formation is quite typical for plant communities. In particular, the similar structures have been observed in Ural Mountains [7].

Historically, there are two approaches to the consideration of spatially distributed systems of plants. The first approach is a phenomenological modeling. For example, in the paper [8] it is proposed reaction-diffusion model of the dynamics of vegetative plants, which describes the pattern formation even in homogeneous and isotropic environments. It includes an equation for the concentration of biomass with diffusion and specially chosen "reactive" term responsible for the plant–plant interaction. Thus, the cause of pattern formation here is the collective interaction of plants.

Another direction in modeling of the forest evolution is an approach based on cellular automata [9–13]. This approach has gone through several stages of development. One of early models was the JABOWA forest model [9] developed for northern forests of New Hampshire. The purpose of JABOWA was to model succession in mixed-species forests and thereby to predict species composition. It was based on the notion that the interactions that drive forest dynamics are local. This model was later modified and extended, for example in [10], to forests ranging from boreal regions to the tropics (FORET model). The next advance was made in [11, 12]. The authors suggested that the range of dynamic behavior that JABOWA-FORET models are capable of exhibiting is too large, since some critical features of the models were simply unknown at that time. They have proposed their own model SORTIE based on that their submodels were designed simultaneously with maximum likelihood estimators necessary to estimate them from simple field measurements [11–13]. SORTIE gave rise to a full pedigree of related models and relevant publications. Probably it is one of the most successful forest simulation models ever developed.

Currently such models include, typically, the individual-oriented behavior. An individual-based model is a class of computational models for simulating the actions and interactions of autonomous individuals (agents) with a view to assessing their effects on the system as a whole. This approach gives a number of advantages such as transparency in relation to the objective biological mechanisms, the ability to describe the system with a high degree of detail to extract more useful information from the simulation results.

It should be noted that the developers of SORTIE, as biologists were intended to describe the evolution of the forest community as accurately as possible. For example, one of the recent variations of the model examines the dynamics of the 12 species [13]. The study of forest pattern formation as the fundamental feature of the spatially distributed system was not in their plans. Therefore from this standpoint we prefer the approach developed in [8]. Unfortunately the model of the forest pattern formation proposed there is phenomenological by its nature.

The main goal of this work is to develop a simple, but still more or less realistic individual-base model in order to study possible pattern formation. We propose a three-component model of the evolution of the taiga boreal forest consisting of the

main tree species of the Perm region and explore the features of its evolution and pattern formation. Our model is based on the submodels of different versions of SORTIE [11–14] but in a highly simplified form.

2 Three-Component Evolutionary Model of Boreal Forest

The vast majority of Russian forests are in the boreal zone. According to the forest inventory data for the Perm region, the forest society consists of a fir (55 % of the total biomass), a birch (26 %) and a pine (12 %). Influence of other trees is not so significant. That is why we proposed a basic model, which is formulated only for these edificators of the ecological system.

In deriving the model, we take into account the following factors:

- Sizes of plants. The tallest tree is a pine—up to 40 meters in height. Birch grows to 20–25 m. Fir may grow up to 15–20 m.
- The fight for the light. Higher trees such as pine, have a competitive advantage. Obscuring his crown young trees, they hamper their growth. However, the need for sunlight in different species of trees is different. It levels the playing field somewhat. A birch is the most sensitive to light, and a fir on the contrary likes to grow in the shade. Thus, a fir easy grows in the shade of the birch, but when it grows, it blocks the growth opportunities for the former. Pine is difficult to rise in the shade, but if it has risen, because of its height it apart from the competition.
- The lifetime. Birch is the most short-lived—60–80 years. Fir and pine lives up to 250 and 400 years respectively.
- Distribution of seeds. Pine has the largest radius of the distribution, but the survival rate of seeds is lowest.

The model consists of three populations of trees and landscape. For simplicity, we consider the area of a square measuring 500 by 500 m. To compute the luminous flux falling on a certain point the landscape, we introduce a uniform grid $N \times N$, where N in the most of the numerical experiments was 100. Populations consist of the sets of individual plants, each of which lives its own life. Wood in the model consists of two cylinders, one of which simulates the crown, and the other – the barrel. Each tree has a number of important current values of its states, first of all this is the height of the whole plant H and the thickness of its trunk D. These variables are tightly related as follows:

$$H = H_1 \left(1 - \exp\left(-\frac{H_2}{H_1} D\right)\right), \tag{1}$$

where $H_{1,2}$ are parameters depending on the type of population. The width and height of the crown is also related to each other

$$R = C_1 d^a, \tag{2}$$

where C_1 and a are the parameters of the populations. Thus, each tree is characterized by a single independent variable. Further, it is assumed that the plant evolves through three stages: (i) seed, (ii) young tree and (iii) mature tree. Seed does not react on light and cannot grow. However, it can take root at a certain place of landscape and become a young tree. The difference between young and mature plants consists in the values of the parameters that determine the rate of growth and the ability to insemination. Each plant involved in evolution at each time step grows in accordance with its status, makes insemination terrain, and may die suddenly.

The model assumes that the main resource for which competing trees, is sunlight. That light is the reason for the growth of plants. The light beam irradiating on each element of the landscape can meet the crown of the tree. The penetration rate λ is the characteristic parameter for each population. If we now integrate over the whole set of plants, one can get the amount of light per unit area of the landscape. After finding the distribution of area illumination for each tree one can calculate its rate of growth for this year. The relative increase in the size of the tree growing on the element of landscape and receiving the light is determined by the Michaelis-Menton law [12]:

$$G = \frac{g_1 g_2 F_{ij}}{2(g_1 + g_2 F_{ij})} d, \tag{3}$$

where $g_{1,2}$ are parameters of populations. Then we can calculate the size of the tree in the next year:

$$D_{t+1} = D_t + G, \tag{4}$$

The next step is the calculation of the death probability of the tree. The probability depends on how fast this tree grows: for good growing trees it is less:

$$M = m_1 \exp(-m_2 G), \tag{5}$$

where m_1 is the mortality rate at zero growth, m_2 is the mortality due to light. If the tree dies, it is permanently deleted from the population [14]. If the tree dies, it is permanently deleted from the population.

The second important process in the model is the production of seeds by mature trees and seeds distribution across the landscape. The insemination is a long-range interaction. Namely this process distributed in space forms a non-linear relationship between the members of the population and contributes to the emergence of non-local structures. For the probabilistic description of the distribution of seeds, we have used the Weibull distribution [11]:

$$V_i = \gamma \sum_{k=1}^{g} \left(\frac{D_k}{30} \right)^2 \exp\left(-r_1 M_{ik}^3 \right). \tag{6}$$

Here D_k is the diameter of the trunk of parent trees, r_1 is the population parameter, and M_{ik} is the distance from i to the k-th parent tree. In our model, we neglect

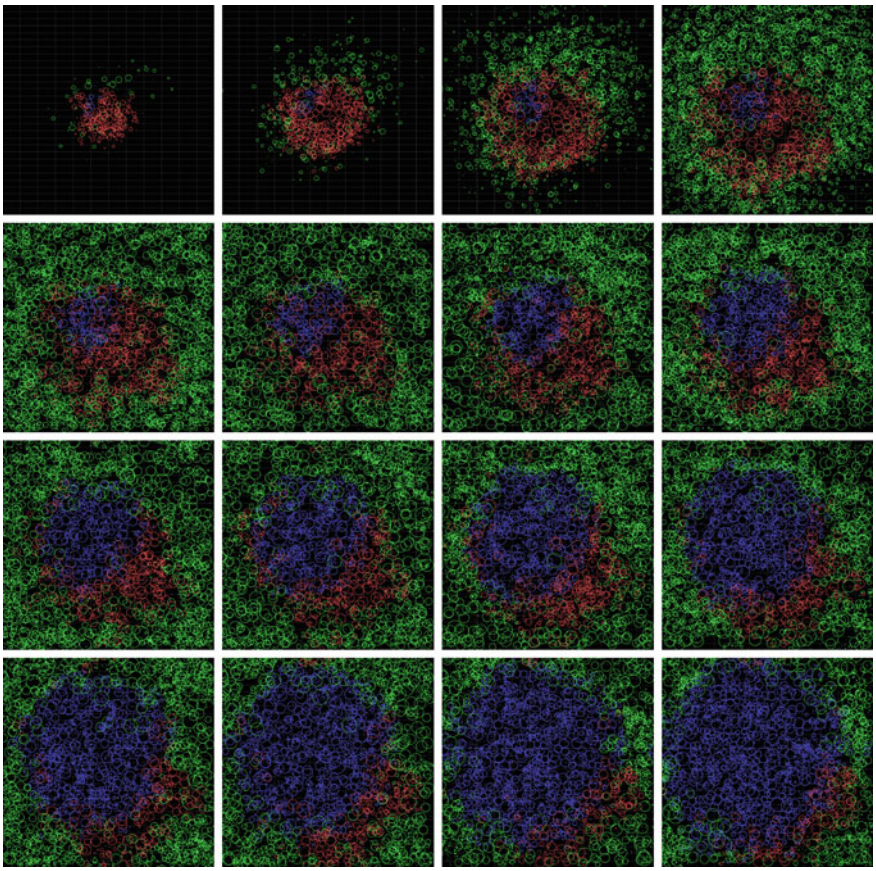


Fig. 1 Evolution of three-component forest from non-uniform initial conditions

the anisotropy of the distribution of seeds (for example, due to the wind) and other possible complicating factors.

The values of parameters calibrated on the basis of available forest inventory data are given in the Table 1. The Java version of on-line forest simulator based on three-component model described above can be found on our website (http://urales.trajan.ru/).

As an example, let us consider the spatio-temporal dynamics of forest simulated within the model let us consider the evolution of the system from an initial state in which the center of empty landscape there are several trees of different species. Figure 1 shows the consistent frames of forest evolution after every 70 iterations (70 years). Complete evolution of the system is equal to 1120 years. Each circle represents a single tree, the width of the circle is proportional to the width of the crown.

Table 1 Parameter values for the three-component model

Type of tree	λ	g_1	g_2	m_1	m_2	r_1
Birch	0.4	0.4	0.05	0.5	2.0	3×10^{-4}
Fir	0.064	0.15	0.15	0.077	6.0	6.9×10^{-4}
Pine	0.4	0.18	0.019	0.268	4.0	1.03×10^{-5}

It is clearly seen that the overgrown of the wasteland occurs due to the birch trees (indicated by red) and pines (green). Firs spread slowly and is only after other species of trees. However, finally the firs displace other trees. Analysis of the distribution of the landscape illumination shows that in the central region occupied by the firs, the light to the earth's surface does not get. This leads to the fact that the seeds cannot take root in this area and fir is slowly but surely expanding their habitat areolas. The very young fir loves twilight, and her living conditions are comfortable.

3 Conclusions

The simple portable individual-based model representing a community of the three populations of forest (fir-birch-pine) has been proposed. Modeling of forest growth, based on the dynamics of individual trees that come into competition with each other for solar energy has shown that the interaction at the micro level lead to the emergence of spatially distributed macroscopic structures.

The work was supported by the Department of Science and Education of Perm region (project C26/244) and Perm State Pedagogical University (project 031-F).

References

1. Murray, J.D.: Mathematical Biology I: An Introduction. Springer, Berlin (2002)
2. Levin, S.A.: The problem of pattern and scale in ecology. Ecology **73**, 1943–1967 (1992)
3. Greig-Smith, P.: Pattern in vegetation. J. Ecol. **67**, 755–779 (1979)
4. Ponce, V.M., Cunha, C.N.: Vegetated earth mounds in tropical savannas of central Brazil: a synthesis. J. Biogeogr. **20**, 219–225 (1993)
5. Kershaw, K.A.: Pattern in Vegetation and Its Causality. Ecology **44**, 377–388 (1963)
6. Valentin, C., D'Herbes, J.M., Poesen, J.: Soil and water components of banded vegetation patterns. Catena **37**, 1–24 (1999)
7. Belkovskaya, T.P., Bezgodov, A.G., Ovesnov, S.A.: Vascular plants of Vishera Reserve. Perm State University, Perm (2004) (Russin)
8. Lefever, R., Lejeune, O.: On the origin of tiger bush. Bull. Math. Biol. **59**, 263–294 (1997)
9. Botkin, D.B., Janak, J.F., Wallis, J.R.: Some ecological consequences of computer model of forest growth. J. Ecol. **60**, 101–116 (1972)
10. Shugart, H.H.: A Theory of Forest Dynamics. Springer-Verlag, New York (1984)
11. Pacala, S.W., Canham, C.D., Silander, J.A.: Forest models defined by field measurements: I. The design of a northeastern forest simulator. Can. J. For. Res. **23**, 1980–1988 (1993)

12. Pacala, S.W., Canham, C.D., Saponara, J., Silander Jr, J.A., Kobe, R.K., Ribbens, E.: Forest models defined by field measurements: II. Estimation, error analysis and dynamics. Ecol. Monogr. **66**, 1–43 (1996)
13. Canham, C.D., Thompson, J., Zimmerman, J.K., Uriarte, M.: Variation in susceptibility to hurricane damage as a function of storm intensity in Puerto Rican tree species. Biotropica **42**(1), 87–94 (2010)
14. Kobe, R.K., Pacala, S.W., Silander Jr, J.A., Canham, C.D.: Juvenile tree survivorship as a component of shade tolerance. Ecol. Appl. **5**, 517–532 (1995)

Part III
Systemic Networking

The Network of the International Criminal Court Decisions as a Complex System

Fabien Tarissan and Raphaëlle Nollez-Goldbach

Abstract Many real-world networks lend themselves to the use of graphs for analysing and modeling their structure. This approach has proved to be very useful for a wide variety of networks stemming from very different fields. Yet, only few papers focused their attention on legal networks. This paper intends precisely to remedy this situation by analysing a major legal network by means of complex system methods. The network under investigation is the network composed by decisions taken by the International Criminal Court since its creation. We first model the network by a simple directed graph in which nodes are the decisions and links represent citations between decisions. Our analysis shows that standard properties shared by common real networks are also present in this network. Then we turn to studying the network by means of bipartite graphs that involve both decisions and articles of law. We show that this two-level structure presents several non trivial properties and we show evidences of the relevance of the bipartite representation to explain properties observed in the graph of citations.

Keywords Complex networks · Bipartite graphs · Legal studies

1 Introduction

Many real-world networks lend themselves to the use of graphs for analysing and modeling their structure. We can cite among others actor networks [1, 2] which relate actors performing in the same movies or authoring networks [2, 3] which relate authors publishing together. Since the seminal paper of Watts and Strogatz [1], it has

F. Tarissan (✉)
Laboratoire d'Informatique de Paris 6, CNRS, Université Pierre et Marie Curie, Paris, France
e-mail: fabien.tarissan@lip6.fr

R. Nollez-Goldbach
Centre de Théorie et d'Analyse du Droit, Université Paris Ouest, ENS, CNRS, Nanterre, France

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_25,

been shown that different kind of networks yet share similar non trivial properties, such as an heterogeneous distribution of their degrees, a high local density, short distances, etc.

Since then this approach has been widely used in many different fields, ranging from computer science (the Internet, peer-to-peer networks, the web), to biology (protein-protein interaction networks, gene regulation networks), social science (friendship networks, collaboration networks), linguistics, economy, etc.Thus, the complex system theme has opened a new interdisciplinary interplay between fields that were usually separated. Recently, law and computer science followed this promising approach, using methods related to complex systems to model and analyse legal networks [4, 5]. It is this rich and promising field of research this paper intends to commit itself to by studying the networks composed of the decisions taken by the International Criminal Court since its creation.

Following the traditional approaches, we use standard directed graphs to represent the networks. Here the nodes stand for the decisions and the links for citations between decisions. Although useful, such a simple representation is not particularly close to the real structure of the networks and does not account for the inherent complexity and hierarchy commonly observed in real data. In the context of legal networks in particular, interactions between decisions take place at various levels. Indeed, to motivate their decisions, judges usually rely on former decisions—first level exhibiting direct interactions—but also refer to articles of the ICC Statute and norms and regulations of the Court—second level pointing out indirect interactions. This induces a two-level structure in which direct and indirect relations interplay in the ruling process.

The existence of two-level structures in real networks led the community to use also *bipartite graphs*, i.e. graphs in which nodes can be divided into two disjoint sets, $\top$ (e.g. articles) and $\bot$ (e.g. decisions), such that every link connects a node in $\top$ to one in $\bot$. Bipartite graphs are fundamental objects which have proven to be very efficient for both the analysis [6–8] and the modeling [9, 10] of complex networks as they are able to reveal patterns that could not have been detected on simple graphs. The present study follows this approach and relies at the same time on simple direct graphs and bipartite graphs in order to extract all relevant properties from the network.

The remaining of the paper is organised as follow: Sect. 2 will review the technical background necessary for going throughout the paper; Sect. 3 will present the main results obtained and finally Sect. 4 will conclude the paper and open on new perspectives.

2 Background

In this section, we introduce the required background for the remainder of the paper. First, we focus on the dataset (Sect. 2.1) extracted from the ICC database. Then, we discuss the several frameworks we used in order analyse its structure (Sect. 2.2).

2.1 Legal Networks

The International Criminal Court (referred to further as ICC) is the first permanent international criminal jurisdiction, established to judge international crimes (genocides, crimes against humanity and war crimes). The creation of the Court is very recent (2002) and only 18 cases are currently opened.

The main production of the Court are decisions, which are legal statement ruling on juridical issues. In order to motivate their decisions, judge may either rely on former decisions of the Court, or rely on articles of the ICC Statute, and/or norms and regulations of the Court. Here below is an example of such a motivation found on a footnote of decision ICC- 01/04- 01/06- 2126- Anx:

> Appeals Chamber, Judgement on the appeals of The Prosecutor and The Defence against Trial Chamber I's Decision on Victims' Participation of 18 January 2008, 11 July 2008, **ICC-01/04-01/06-1432** , para. 95. See also Trial Chamber II, Decision on the Modalities of Victim Participation at Trial, 22 January 2010, **ICC-01/04-1/07-1788**, para. 30. See also Defence for Germain Katanga's Additional Observations on Victims' Participation and scope thereof", 10 November 2009, **ICC-01/04-01/07-1618**: "It has been held that **article 69(3)** gives the Court a general ...

In this example, one can notice the two types of arguments used by the judges. The text clearly refers to former decisions (highlighted in red) but also refers to article of the ICC Statute (in blue).

In the rest of the paper, we only focus on the Lubanga case (DRC situation) to concentrate on the most advance case. It involves approximately 7,000 decisions and 1,500 articles.

2.2 Graph Frameworks

Directed graphs. As depicted in the introduction, it is quite natural to represent the network as a directed graph $G = (V, E)$, with $n = |V|$ and $m = |E|$, in which the nodes stand for the decisions – identified by their ICC number – and a link between nodes u and v exists if the decision u cites the decision v. Note that technically, the graph is a Directed Acyclic Graph (DAG) since, for obvious reasons, the decisions can only refer to existing ones. Thus there is no cycle in the network.

This defines the graph of citations among decisions, which will be referred to simply as the *graph of citations*. By doing so, we can compute standard metrics and compare the results to what is obtained on usual complex networks. According to standard studies, one usually observe for instance that graphs are sparse, *i.e.* the density $\delta = \frac{2.m}{n.(n-1)}$ is very small, and their degree distribution is heterogeneous (often close to a power-law).

Another key property concerns the local density which is meant to study how dense a neighbourhood of a node is in the graph. This concept is generally captured

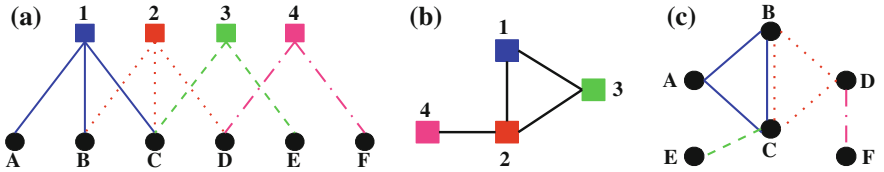


Fig. 1 Example of bipartite graph and its $\{\top, \bot\}$-projections

by the clustering coefficient $\mathtt{cc}(G)$ or the transitivity ratio $\mathtt{tr}(G)$ [1, 11, 12], defined formally by:

$$\mathtt{cc}(G) = \frac{\sum_v \frac{\Delta(v)}{\vee(v)}}{n} \qquad \mathtt{tr}(G) = \frac{\Delta(G)}{\vee(G)}$$

where, for each $v \in V$, $\Delta(v)$ denotes the number of directed triangles (sets of three nodes u, v, w, such that $(u, v), (u, w), (v, w) \in E$) which originate at v; $\vee(v) = \frac{d(v).(d(v)-1)}{2}$ denotes the number of pairs of neighbours of v which computes the number of *possible* directed triangle; $\Delta(G) = \sum_v \Delta(v)$; and $\vee(G) = \sum_v \vee(v)$. Note that the clustering coefficient of a node can be defined for the in-degree and the out-degree, that is when the node is at the origin of the directed triangles (case of u in the example above) and when it is at the end (case of w). Both variants makes sense and will be investigated in Sect. 3.

A classical observation in complex network studies is that all these quantities are high, at least compared to the density δ of the graph. Note however that the meaning of the existence of such a pattern depends on the context of the network. It has been shown that it could be related to robustness properties of the network, or properties related to dynamical aspects of the networks (see for instance [13, 14] for biological cases).

Bipartite graphs. As stated in the introduction, the previous formalism does not account for higher level of relations between the decisions. In particular, one does not exploit the references made to the articles which relate the decisions to the articles (and regulations and norms) they refer to. This two-level structure calls for a specific framework that is perfectly matched by the concept of bipartite graphs.

A *bipartite graph* is a triplet $G_b = (\top, \bot, E_b)$, where $\top$ is the set of *top* nodes (here the articles/norms/regulations), $\bot$ the set of *bottom* nodes (here the decisions), and $E_b \subseteq \top \times \bot$ the set of links that relate the decisions to the articles. We denote by $n_\top$ (resp. $n_\bot$) the number of top nodes (resp. bottom nodes).

Compared to standard graphs, nodes in a bipartite graph are in two disjoint sets, and the links are always between a node in one set and a node in the other set. An example of bipartite graph is given in Fig. 1a, where $\top$ nodes are depicted by squares and $\bot$ nodes by circles.

The $\bot$-projection of G_b is the graph $G_\bot = (\bot, E_\bot)$ where two nodes (of $\bot$) are linked together if they have at least one neighbour in common (in $\top$) in G_v: $E_\bot = \{(u, v), \exists x \in \top : (u, x) \in E_b \text{ and } (v, x) \in E_b\}$. The $\top$-projection is defined dually.

Both projections are illustrated in Fig. 1b and c. Thus, in our case, the $\perp$-projection corresponds to a graph of decisions, such as G, but a link between two decisions exists if and only if there is at least one common article to which they both refer.

Note that by projecting a bipartite graph into a simple graph, we can then reuse all the metrics defined above for standard graph. But we can also compute specific metrics for bipartite graphs, such as $k_\top$ (resp. $k_\perp$) the average degree of top nodes (resp. bottom nodes) and $\delta_\flat = \frac{m_\flat}{n_\top . n_\perp}$ the density of the bipartite graph.

3 Results

The purpose of this section is to position the ICC decisions network as regard the properties observed in common complex networks. To do so, we start investigating global statistics (Sect. 3.1) before focusing on more specific properties such as the degree distribution, the local density and some correlations in the bipartite structure (Sect. 3.2).

3.1 A Global Perspective

The first statistics we focus on concern some basic properties observed in most real-world networks, formally presented in the previous section. Table 1 presents the results both for the graph of citations (left) and the bipartite graph (right).

As expected, all usual observations made on real-world networks stand also for the legal network under investigation. More precisely, one can see that the graph is sparse ($\delta = 7 \times 10^{-4}$) and that the local density (both the transitive ratio and the clustering coefficient) are several orders of magnitude higher.

Beside one can notice that the highest degree d^+ is also way higher than the average degree k. This indicates some heterogeneity in the degree distribution that will be investigated further.

As regard the statistics on the bipartite graphs, it also matches the expectations since the graph is also sparse and both top and bottom highest degrees ($d^+_\top$ and $d^+_\perp$) are two orders of magnitude higher than their respective average degree ($k_\top$ and $k_\perp$).

Those results on the global structures of the two graphs are now refine in the following section.

3.2 A Deeper Analysis

Degree distribution. In order to refine the general statistics presented above, Fig. 2 presents the degree distribution observed in the graph of citations (left) and the bipartite graph (right). In the two figures, the horizontal axis, in log-scale, stands for the degree of the nodes, while the vertical axis, also in log-scale, presents the inverse

Table 1 Global statistics for the graph of citations (left) and decision/article bipartite graph (right)

	Directed graph	
	Complete	Connected Component
n	6 894	1 575
m	17 625	3 319
δ	0.00	0.00
k	2.5	2.1
d^+	214 / 433	158 / 433
tr	0.03 / 0.10	0.03 / 0.08
cc	0.14 / 0.19	0.12 / 0.22
	Bipartite graph	
$n_\top$	1 415	
$n_\bot$	6 894	
$m_\mathfrak{b}$	11 371	
$n_\bot(conn.)$	1 683	
$k_\top$	1.7	
$k_\bot$	8.0	
$\delta_\mathfrak{b}$	0.00	
$d^+_\top$	802	
$d^+_\bot$	116	

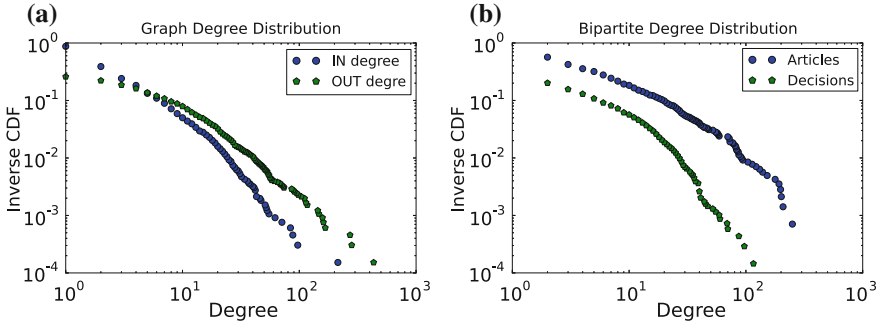


Fig. 2 Inverse cumulative degree distributions for the graph of citations (*left*) and the bipartite graph (*right*)

cumulative proportion (according to the total number of nodes). The shape of the plots, close to a straight line in the log-log scale over several orders of magnitude, thus confirms that we are dealing with heterogeneous distributions.

One can notice that for the out degree (the number of citations a decisions makes) in the graph of citations, the proportion is higher than the in-degrees (the number of references made to a decision). This is in particular true for the high degrees, which is well explained by the existence of annexes in the corpus. Indeed, those usually list all the decisions that a case has referred to.

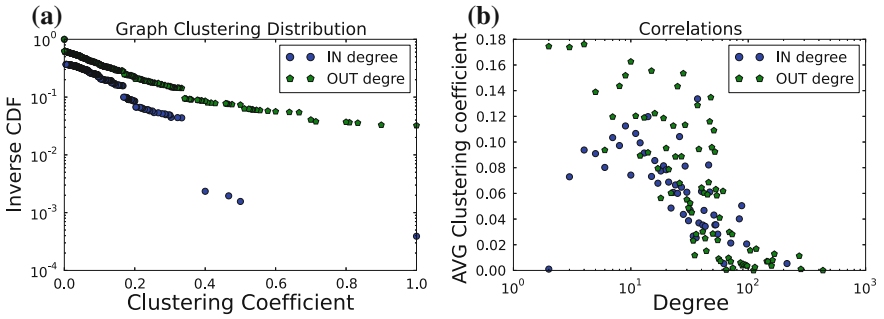


Fig. 3 Clustering coefficient and correlations in the graph of citations

Note also that, although the high out-degrees are not particularly meaningful since it corresponds to annexes, it is however particularly relevant for the in-degrees. Indeed, a manual investigation showed that the top-3 of cited decisions corresponds for major ruling in the case. The first one deals with the conviction of the accused, the second with its sentence and the last with the remedy and reparations of the victims, which are all three important issues in the international Court.

As regard the bipartite structure, one can also notice the same kind of comparison between the degrees of the articles and the ones of the decisions. This indicates that, from a global point of view, each article tends to be more referred to that each single decisions. This can be explained by the status of the articles of the ICC Statute towards decision. This remark is also corroborated by the highest degrees of articles and decisions (see Table 1 right).

Clustering coefficient. Figure 3 presents several properties related to the clustering coefficient, as defined in Sect. 2. The left part displays the inverse cumulative distribution (in lin-log scale) of the coefficients in the graph of citations defined for the nodes that originates the directed triangles (out-degree, in green) and the ones that are at the end of the directed triangles (in-degree, in blue). Note that the plots are normalised over the number of nodes with degree ≥ 2 in order to avoid side effects from the nodes of degree 1, for which the notion of clustering coefficient is inadequate.

Interestingly, one can notice that the proportions of nodes that are at the origin of triangles is particularly high compared to the one that are at the end. This indicates that when a decision refers to two other decisions, one of them tends to cite the last one very often. Manual investigations performed on the triangles that involve highly cited decisions showed that the over-representation of this pattern is meaningful as regard legal networks. It turned out that most of those triangles involve a decision w that is contested by the accused, thus leading to a decision v that cite w. Then a final decision u rules between the issue, thus citing both v and w. This leads to the directed triangle originated by u with w as end-point.

The right part of Fig. 3 presents a non trivial correlation between the clustering coefficient of a node and its degree in the graph of citations. More precisely, a dot

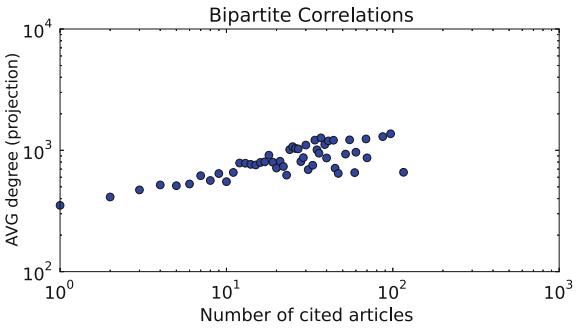


Fig. 4 Correlations

(x, y) in this plot corresponds to the fact that nodes with degree x have, in average, a clustering coefficient of y. The figure shows that the higher the degree, the lower the clustering coefficient in average. This is true both for in-degree and out-degree except for very low degrees. Indeed, the case of degree-2 nodes is very different depending on whether we consider the origin of the triangles or the end. For the end, the coefficient is very low (close to 0), which shows that when two decisions cite a third one, they usually do not rely on each other (blue dots). On the other end, if one decision cite two different decisions, those two decisions tends to be related (green dots). The latter case is easily explained by the remark made in the previous paragraph since in the former example, the decision u has precisely degree 2 and its clustering coefficient is 1 thus increasing the average clustering coefficient of degree-2 nodes.

Bipartite/Projection correlations. Finally, Fig. 4 presents a correlation between the degree of a decision in the bipartite graph and its degree in the projection. More precisely, a dot (x, y) in the plot stands for the fact that decisions that refer to x articles in the bipartite network are related (in the projection) to y different decisions in average.

The figure presents a strong correlation between those two quantities. The regular increase (in log-log scale) is natural since the more a decision refers to different articles, the more decisions it will be related to. What is less intuitive though, is why the slope of the progression is not higher. Indeed, the reader might have noticed that according to the plot, multiplying by 100 the number of references to articles only multiplies the degree in the projection by 10. This fact clearly indicates a strong overlapping between the articles. This observation makes sense since decisions that concerns a similar content (that is, deals with similar legal issues) tend to refer to the same articles to motivate their content.

4 Conclusion

In this paper, we studied a legal network composed of decisions taken by the International Criminal Court since its creation. The analysis made on the most advance case of the Court shows that it completely matches standards properties shared by real-world networks, thus confirming the relevance of the complex system approach towards legal networks.

Besides, we investigated more in depth some of the properties related to the local density and provided a first interpretation of the over-representation of directed triangles in such networks. By using different formal frameworks, we also exhibited new properties such as the overlapping of articles in the bipartite representation of the decision/article network.

However, the different analyses performed in this study have been made independently. Yet, the strong patterns identified here call for the definition of a unique framework able to integrate both direct relations between decisions (such as the citation process) and indirect relations (such as the decision-article relations). This would entitle to consider correlations between those two levels of interaction that might shed light on important properties of the network. We let this promising approach for further studies.

Another appealing aspect which has not been investigated for the moment concerns the time. As stated in Sect. 2, the graph is actually a DAG since temporal aspects prevent from the existence of cycles in the network. This remark lead to consider the temporal evolution of the network instead of considering the whole decision network since the creation of the Court. We also let this aspect for future studies.

Acknowledgments This work is partly funded by the National Center for Scientific Research (CNRS) through the PEPS Project "DoRé".

References

1. Watts, D., Strogatz, S.: Collective dynamics of small-world networks. Nature **393**, 440–442 (June 1998)
2. Newman, M., Strogatz, S., Watts, D.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64 (2001)**
3. Newman, M., Watts, D., Strogatz, S.: Random graph models of social networks. PNAS **99**, 2566–2572 (2002)
4. Bourcier, D., Mazzega, P.: Codification law article and graphs. In: Lodder, A.R., Mommers, L. (eds.) Legal Knowledge and Information Systems, JURIX 2007. IOS Press, Amsterdam 29–38 (2007)
5. Boulet, R., Mazzega, P., Bourcier, D.: A Network Approach to the French System of Legal codes - Part I: Analysis of a Dense Network, Artificial Intelligence and Law, 19, 333–355, 2011.x
6. Latapy, M., Magnien, C., Del Vecchio, N.: Basic notions for the analysis of large two-mode networks. Soc. Netw. **30**(1), 31–48 (January 2008)
7. Ahn, Y.-Y., Ahnert, S.E., Bagrow, J.P., Barabási, A.-L.:Flavor network and the principles of food pairing. Nature Scientific Reports(2011)

8. Tumminello, M., Micciché, S., Lillo, F., Piilo, J., Mantegna, R.: Statistically Validated networks in bipartite complex systems. PLoS ONE, **6 (2011)**
9. Guillaume, J.-L., Latapy, M.: Bipartite graphs as models of complex networks. Physica A **371**, 795–813 (2004)
10. Tarissan, F., Quoitin, B., Mérindol, P., Donnet, B., Latapy, M., Pansiot, J-J.: Towards a bipartite graph modeling of the internet topology. Comput. Netw. **57(11), 2331–2347 (2013)**
11. Schank, T., Wagner, D.: Finding, counting and listing all triangles in large graphs, an experimental study. In: Workshop on Experimental and Efficient Algorithms (WEA) (2005)
12. Schank, T., Wagner, D.: Approximating clustering coefficient and transitivity. J. Graph Algorithms Appl. (JGAA) **9**(2), 265–275 (2005)
13. Mangan, S., Itzkovitz, S., Zaslaver, A., Alon, U.: The incoherent feed-forward loop accelerates the response-time of the gal system of Escherichia coli. JMB **356**, 1073–1081 (2006)
14. Mangan, S., Alon, U.: Structure and function of the feed-forward loop network motif. PNAS **100**, 11980–11985 (2003)

Inference of Optimized Control Strategies for Genetic Networks

Natalja Strelkowa

Abstract In this chapter we present the application of control theoretical concepts to stochastic dynamical systems which are based on the current knowledge of genetic networks. We showcase the application of reinforcement learning algorithm inferring an optimized control strategy for a genetic switch reversal. The approach does not require precise knowledge of gene network mathematical equations and is therefore also applicable to experimentally obtained time traces.

Keywords Feedback control · Genetic switch

1 Introduction

Even though natural biological systems such as gene regulatory networks might appear to be very different from classical engineering examples as for instance flight control or nuclear reactor shutdown, the system regulation is believed to be very similar [1]. In both cases robust functionality under changing environments is achieved via feedback control loops and can be explained with the concepts of control theory (see Fig. 1).

We consider the application of feedback control to gene regulatory networks. Perturbations or intervention signals for genetic networks can be implemented biochemically via conditional gene knock outs [2, 3], heat shocks [4], light pulses [5, 6], etc. Application of control theoretical concepts would provide intervention protocols for these biochemical perturbations, which would direct genetic networks to desired states in an optimized way by accumulation and integration of information from previous experiments.

N. Strelkowa (✉)
Boehringer Ingelheim Pharma GmbH and Co. KG, Rhineland-Palatinate, Germany
e-mail: natalja.strelkowa@gmail.com

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_26,

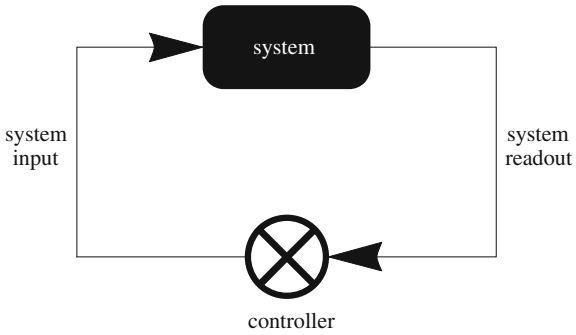


Fig. 1 Classical idea of feedback control. The current state of the system or readout is measured and used as input into the controller. Based on the readout, the controller will infer appropriate interference action and an input signal will be feed back into the system

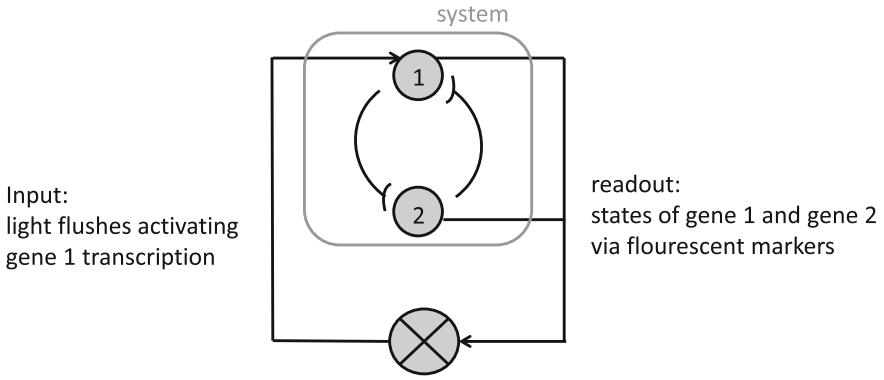
The optimization criteria can be defined in a system tailored manner. It can for example be time, i.e. use the control signals to reach the desired state as quickly as possible, or cell burden, i.e. use the control signals to reach the desired state minimizing the expression of heterologous proteins, or a weighted sum of both factors.

Reinforcement-learning can be used for deduction of optimized control strategies [7]. This algorithm uses time traces to induce a multi-dimensional function $Q(\mathbf{n}, \mathbf{u})$, which is the score of a control action $\mathbf{u} \in U$ for a system state $\mathbf{n} \in \mathbb{N}^s$. The optimized intervention strategy for direction of the network to a desired state is obtained by taking the maximum of this score function at each state $\mathbf{n}$: $\max_{\mathbf{u}\in U} Q(\mathbf{n}, \mathbf{u})$. The algorithm induces optimized control strategies based on time series obtained from the dynamical system, the definition of the desired state and the cost function. It is not dependent on a specific form of the model, i.e. largely model independent.

We demonstrate the application of the algorithm on a genetic network: toggling a noisy genetic switch.

1.1 Controlled Reversal of a Genetic Switch

Switches are typical and important constructs for the forward designed [8] as well as for the naturally occurring circuits. One example, where cascades of switches occur and are believed to dominate the dynamics are antagonistic switches in developmental biology [9, 10]. The cell fate is believed to be determined by successive switch reversals of the phenotype dominating genes. For the stem cells tissue engineering intervention protocols are desired, which would direct the cells into a particular fate both fast and without protein over-expression, which reduces the fitness of the cells. An optimized strategy for switch reversals avoiding over-expression can be sought using the reinforcement learning algorithm.



We use a simple design for a switch to demonstrate the formulation of this biological problem in control theoretical terms. A classical topology for a biological switch are two genes mutually repressing each other (figure above). If the repression is strong enough then this system will show bi-stability [11, 12], which means in the steady state either the first gene is highly expressed and the second gene is suppressed ($p_1 \uparrow$, $p_2 \downarrow$) or vice versa ($p_2 \uparrow$, $p_1 \downarrow$).

Stochastic dynamics of the genetic switch can be formulated as a Master equation based on the basic mass-action kinetics and using time scale separation between the slow protein and fast mRNA dynamics (see for instance [13] for Master equation derivation) $P(p_1 = n, p_2 = m; t) \equiv P(n, m; t)$:

$$\begin{aligned} d_t P(n,m;t) \; &= (g(m) + q_s u) P(n-1,m;t) - (g(m) + q_s u) P(n,m;t) \\ &+ g(n) P(n,m-1;t) - g(n) P(n,m;t) \\ &- d_1 n P(n,m;t) + d_1 (n-1) P(n-1,m;t) \\ &- d_2 m P(n,m;t) + d_2 (m-1) P(n,m-1;t) \end{aligned} \tag{1}$$

where $g(x) = \frac{e}{1+x^2}$ is the Monod-type repression term. The order of magnitude for the constants is based on [14, 15] and to illustrate the procedure we assume a symmetric switch $e = 400; d_1 = d_2 = 1$. The number of proteins per unit time produced as the result of the light control signal is denoted by q_s and u as usually denotes the control actions $u \in \{0, 1\}$.

Control problem formulation for a genetic switch reversal. We assume that for both genes the protein concentrations are given as readouts for instance via fluorescent markers. The initial condition is the vicinity of the fixed point ($p_1 \downarrow$, $p_2 \uparrow$), where gene 1 is silent and gene 2 is expressed. The goal of the control strategy is to direct the switch to the second fixed point ($p_1 \uparrow$, $p_2 \downarrow$) by applying control signals to the first gene. The control signals increase the concentration of the first gene in small increments, which biochemically corresponds to light flashes activating a photosensitive promoter. Biochemically the response occurs on fast time scales (order of

seconds [5]) compared to the gene expression time scales occurring on the order of minutes or even hours and therefore we can approximate the action of the light inducer as a discrete set light on and light off $U = \{0, 1\}$. In classical control theory this type of control is often referred to as "bang bang".

To create the training set we start from random initial conditions and the "light on" signals $u(t) = 1$ are applied with the probability 3 % covering the interval of interest $[0, 400] \times [0, 400]$. The instantaneous reward function is defined in the following way: Two boundary conditions are used to specify the minimal and the maximal boundary for the instantaneous reward function. The first absorbing boundary condition is given by the minimal reward -1, which is assigned to the neighborhood of fixed point B($p_1 \downarrow$, $p_2 \uparrow$), which is the initial state we aim to reverse. By absorbing in this context we mean that if the system reaches that state after a series of random signal applications, we assign the according instantaneous reward to the last point and reset the simulation from random initial conditions. The second absorbing boundary with the maximal reward 1 is assigned to the neighborhood of the fixed point ($p_1 \uparrow$, $p_2 \downarrow$), which we are trying to achieve. Each application of the control signal is associated with a cost for the cell to express the foreign protein and therefore the instantanious reward is diminished if the concentration of the first protein is too high, i.e. $\backsim$ exp(-p1). In summary the reward function is:

$$r(\mathbf{p}, u, t) = \begin{cases} -1 & \text{if } \mathbf{p} \in B(p_1 \downarrow, p_2 \uparrow) \\ 1 & \text{if } \mathbf{p} \in B(p_1 \uparrow, p_2 \downarrow) \\ \text{exp(-p1/e)} & \text{else} \end{cases} \tag{2}$$

Algorithmic outcomes and interpretation. The intuitive outcome for a single switch is a threshold rule: if the concentration of the up gene is low then apply induction signal. However, the main disadvantage of this technique is that it does not take into account the system dynamics and the threshold value is not clear a priori. In the pioneering synthetic biology implementation by Gardner et al. [14] the chemical inducer IPTG has been applied during the whole switching time of 6 hours to the bacterial population. This technique can not be applied if not just one but several switches need to be reversed.

Illustrating our approach we have induced an optimized control strategy for the switch reversal, which should avoid over-expression of the first protein (see Fig. 2). Indeed if starting from initial conditions where the first gene is down and the second is up as shown in the picture only few kicks are required in order to reverse the switch. From the input-output relationships contained in the training time series the reinforcement learning technique deduced a control strategy, which takes advantage of the internal system dynamics.

Implications of noise. One of the challenges in controlling gene expression is that molecular dynamics of the involved biochemical reactions are very noisy [13, 16]. The stochasticity due to low copy number involved in the genetic transcription process is inherent and requires that the proposed control strategy is robust to noisy algorithmic input.

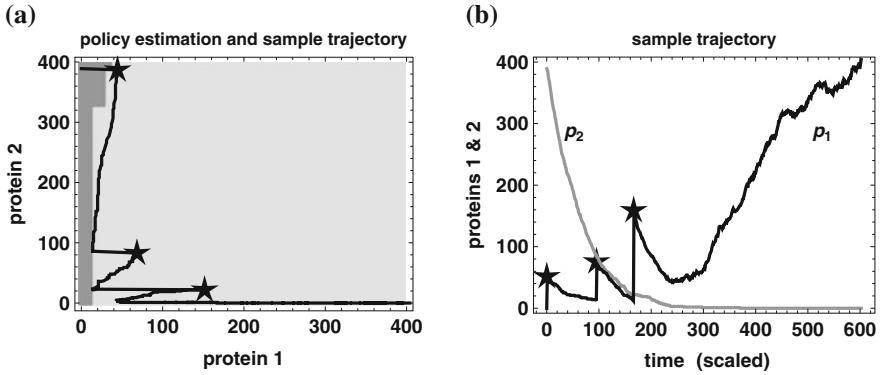


Fig. 2 Optimized reversal of a genetic switch **a** Induced state dependent control policy. The *dark gray area* indicates the set $\{Q(\cdot, 1) > Q(\cdot, 0)\}$, which means genetic network states, where the value for the action "apply control" $u = 1$ is larger than for "do not apply control" for $u = 0$. The complementary set within the boundary conditions is shown in *light gray* **b** Simulated trajectories for $p_1(t)$ (*black*) and $p_2(t)$ (*gray*) starting from B($p_1 \downarrow$, $p_2 \uparrow$) is controlled. Using the algorithmically estimated policy shown in **a** towards the target state ($p_1 \uparrow$, $p_2 \downarrow$). System states for which the control action has been applied ($u = 1$), are marked by *black stars* in both sub-figures. *N*ote: Conducting the switch to the desired state with this algorithmic induced policy avoids over-expression. In comparison to the induced policy, naive or blind application of the control action on the first protein and waiting till the second protein will reach smaller concentration would result in 1000-fold over-expression and therefore lead to unnecessary cell burden

The reinforcement learning algorithm achieves the robustness with respect to noise in gene network dynamics via an embedded regressor EXTremely RAndomized Trees (Extra-Trees) [17].

Our input time traces of the genetic switch dynamics have been simulated using Gillespie algorithm and explicitly taking the inherent stochasticity due to low copy number into account. Several independent algorithmic runs have confirmed that indeed the reinforcement learning algorithm is robust with respect to noisy algorithmic inputs.

2 Discussion

Previously discussed feedback control strategies based on the input-output relationships can now also be applied to genetic networks due to the recent progress in technology for observing and influencing gene regulatory networks. Quantitative *in vivo* estimates of biological system states can for example be obtained via fluorescent markers [18, 19] and spatially targeted transcription induction can be performed in living cells using monochromatic light [5, 6]. With these tools classical engineering concepts of optimal feedback control, where a quantitative system state estimate (read-out) and targeted intervention (control action) are required, is likely to become *in vivo* realizable also for genetic circuits.

References

1. Csete, M.E., Doyle, J.C.: Reverse engineering of biological complexity. Science **295**, 1664–1669 (2002)
2. Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., et al.: Dissecting self-renewal in stem cells with rna interference. Nature **442**, 533–538 (2006)
3. Liu, Y., Asakura, M., Inoue, H., Nakamura, T., Sano, M., et al.: Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. Proc. Natl. Acad. Sci. USA **104**, 3859–3864 (2007)
4. Mettetal, J.T., Muzzey, D., Gomez-Uribe, C., van Oudenaarden, A.: The frequency dependence of Osmo-adaptation in Saccharomyces cerevisiae. Science **319**, 482–484 (2008)
5. Shimizu-Sato, S., Huq, E., Tepperman, J.M., Quail, P.H.: A light-switchable gene promoter system. Nat. Biotech. **20**, 1041–1044 (2002)
6. Levskaya, A., Weiner, O.D., Lim, W.A., Voigt, C.A.: Spatiotemporal control of cell signalling using a light-switchable protein interaction. Nature **461**, 997–1001 (2009)
7. Ernst, D., Geurts, P., Wehenkel, L.: Tree-based batch mode reinforcement learning. J. Mach. Learn. Res. **6**, 503–556 (2005)
8. Andrianantoandro, E., Basu, S., Karig, D.K., Weiss, R.: Synthetic biology: new engineering rules for an emerging discipline. Mol. Sys. Bio. **2**(2006), 0028 (2006)
9. Waddington, C.H.: Organisers and Genes. Cambridge Univ Press, Cambridge, UK (1940)
10. Graf, T., Enver, T.: Forcing cells to change lineages. Nature **462**, 587–594 (2009)
11. Smith, H.: Oscillations and multiple steady states in a cyclic gene model with repression. J. Math. Biol. **25**, 169–190 (1987)
12. Strelkowa, N., Barahona, M.: Switchable genetic oscillator operating in quasi-stable mode. J. R. Soc. Interface **7**, 1071–1082 (2010)
13. Strelkowa, N.: How Nature Works. Springer International Publishing, Switzerland (chapter Stochastic Complexity Analysis in Synthetic Biology. Emergence, Complexity and Computation, Vol. 5) (2013)
14. Gardner, T., Cantor, C.R., Collins, J.J.: Construction of a genetic toggle switch in escherichia coli. Nature **403**, 339–342 (2000)
15. Guantes, R., Poyatos, J.F.: Dynamical principles of two-component genetic oscillators. PLoS Comput. Biol. **2**, e30 (2006)
16. Swain, P.S., Elowitz, M.B., Siggia, E.D.: Intrinsic and extrinsic contributions to stochasticity in gene expression. Proceedings of the National Academy of Sciences of the United States of America, vol. 99, pp. 12795–12800 (2002)
17. Geurts, P., Ernst, D., Wehenkel, L.: Extremely randomized trees. Mach. Learn. **63**, 3–42 (2006)
18. Cai, L., Friedman, N., Xie, X.S.: Stochastic protein expression in individual cells at the single molecule level. Nature **440**, 358–362 (2006)
19. Bennett, M.R., Hasty, J.: Microfluidic devices for measuring gene network dynamics in single cells. Nat. Rev. Genet. **10**, 628–638 (2009)

Network Topologies for Cellular Automata Computation

Camelia Chira and Anca Andreica

Abstract The density classification problem aims to find automata able to correctly classify the density of the initial configuration. This problem is highly challenging as the desired computation requires global coordination while Cellular Automata (CAs) rules rely on the local interaction of simple components. Instead of using the standard CA topology of regular lattice, the current chapter focuses on network topologies that can be used in connection with a simple fixed rule in CA computation. The state of a cell evolves according to the majority of its neighbors in the network. In this chapter, we propose a hill-climbing approach to find good network topologies for the density classification problem starting from initial small-world networks. The network solution space is searched in a random hill-climbing manner based on a simple mutation operator changing the network each iteration. Experiments emphasize the identification of network topologies with a good performance for CA computation. The best identified networks are further studied under a dynamic framework to test their robustness against failures and changes that might occur in the network. Results confirm a good sustained performance of networks identified using hill-climbing search.

Keywords Cellular automata · Density classification task · Complex networks · Hill-climbing search

C. Chira · A. Andreica (✉)
Department of Computer Science, Babes-Bolyai University, 1 Kogalniceanu, 400084 Cluj-Napoca, Romania
e-mail: anca@cs.ubbcluj.ro

C. Chira
e-mail: cchira@cs.ubbcluj.ro

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_27,

1 Introduction

Cellular Automata (CAs) represent important tools in the study of complex systems and interactions. CAs are decentralized structures of simple and locally interacting cells that evolve following a set of rules [21]. The one-dimensional binary-state CA capable of performing computational tasks has been extensively studied in the literature [7, 11, 13, 16]. The problem of finding CA rules able to generate a desired global behavior is highly challenging because only local information influences the cell evolution. The density classification task (DCT) is one such problem. DCT aims to find a binary one-dimensional CA able to classify the density of 1s in the initial configuration. The CA lattice starts with a given binary string called the initial configuration and, after a maximum number of iterations, the CA will reach a certain configuration. If this is formed of homogeneous states of all 1s or 0s, it means that the initial configuration has been classified as density class 1, respectively 0. Otherwise, the CA makes by definition a mis-classification [15]. The performance of a rule measures the classification accuracy of a CA based on the fraction of correct classifications over 10^4 initial configurations selected from an unbiased distribution. It has been shown that there is no rule that can correctly classify all possible initial configurations [8].

The standard CA topology used is a regular lattice and the state of a cell evolves depending on the neighboring cells. Most existing studies focus on developing algorithms able to find high performant rules for one-dimensional CAs with lattice topology. Evolutionary algorithms [6, 9, 11, 13, 14], genetic programming [1], coevolutionary learning [7] and gene expression programming [5] have been shown to be well-suited to detect valuable CA rules. The performance of the best rules detected by genetic algorithms is around 0.76 [6, 9, 15] while coevolutionary learning [7] and multiobjective evolutionary model [12] obtained rules with a performance above 0.8.

In [19], CAs computation on small-world networks is first studied. In this approach, the topological structure of CAs refers to general graphs and cells change states according to neighboring cells in the network. The rule is the same for all cells and the objective of the problem shifts from finding the best CA rule to finding the best CA network topology. In [3, 4, 17], small-world type network topologies are evolved starting from an initial population of regular and random structures. The performance in the case of DCT is above 0.8 for the small-world networks constructed by Watts [19] as well as for the networks evolved by Tomassini et al. [17]. Moreover, a recent study investigating the influence of the CA topology on the evolved dynamics indicates the existence of topology-induced phase transitions and topological bifurcation points [2].

In this chapter, we use a hill-climbing approach to generate the network topology and test its performance for DCT in one-dimensional CAs. Each iteration, the configuration is changed by using a simple mutation operator and the new configuration is accepted to replace the existing one based on a fitness function. Experiments show that this simple approach is effective in constantly finding CA topologies which lead

to a performance above 0.8 for the DCT. The robustness of the obtained small-world networks is evaluated under a dynamic framework as the topology is perturbed by edge removals and/or additions. Results emphasize that perturbations on the network topology do not cause the failure of the system and the performance is not affected by network dynamics.

The rest of the chapter is structured as follows: Sect. 2 describes DCT and briefly presents related work in the area of network based CA computation, Sect. 3 presents the hill-climbing approach to generate network topologies and the main results obtained, Sect. 4 presents the analysis of network dynamics based on several computational experiments, and Sect. 5 contains the conclusions and directions for future research.

2 Cellular Automata Computation on Small-World Networks

Networks represent a central model for the description of complex phenomena and they have been studied in fields such as mathematics, neuroscience, biology, epidemiology, sociology, social psychology and economy. Real-world networks have structural properties that set them appart from regular lattices and random graphs [19]. For instance, small-world networks are characterized by the presence of short paths and a large clustering coefficient. Watts [19] described the "small world effect" property of networks, i.e. the average distance between vertices in a network is short, usually scaling logarithmically with the number of vertices in the network. The clustering coefficient refers to the probability that two vertices that are both neighbors to the same third vertex are also neighbors of each other (also called network transitivity).

Watts and Strogatz [18, 19] studied the computational properties of small-world networks by examining CAs computation on this type of networks. In this approach to DCT, the rule is fixed and the performance of different small-world networks for DCT is evaluated. A small-world graph is constructed starting from a regular ring of nodes in which each node has k neighbors. A random rewiring procedure is then applied as follows [18, 19]: a vertex and the edge connecting it to a neighbor is chosen, and the edge is reconnected with probability p to a vertex uniformly chosen at random from the entire ring. This process is repeated by moving clockwise around the ring until each vertex is considered once (in connection with nearest neighbors, second-nearest neighbors, etc. - depending on the value of k). This way, a number of shortcuts (i.e. edges that link nodes which would be more than two edges apart if they were not directly connected) are produced. Watts and Strogatz [18] observe that for intermediate values of p the graph constructed in this way is a small-world network (with small characteristic path length and high clustering coefficient).

A simple majority rule is used for small-world network DCT defined as follows: at each time step, each node takes the state of the majority of its neighbor nodes in the graph (if the number of state 1s equals the number of state 0s in the neighbors list then the node is randomly assigned a state with equal probability between 0 and

1). Small-world networks have a performance around 0.8 for the DCT with the fixed majority rule for CAs of size 149.

Tomassini et al. [3, 4, 17] investigated network based CAs for the density and synchronization problems. Spatially structured evolutionary algorithms are engaged to find the best performant network topology for DCT when the rule is fixed to the majority rule. An individual represents a network structure and the fitness is computed based on the fraction of initial configurations (out of 100 initial configurations generated anew for each individual) correctly classified by the majority rule based on the neighborhood given by the network. The initial population is generated in two ways: starting from regular rings with node degree $k = 4$ (slightly perturbed by adding a link for each node with a low probability) and random graphs. The best evolved network starting from initial regular rings has a performance of 0.823 (for 149 cells) while the result for random graphs as initial population is similar (performance of 0.821 of the best network). The robustness of the evolved topologies is tested in two scenarios as follows: (i) *probabilistic faults* [17] which allow the rule of each cell to yield an incorrect output state with a certain probability (the structure of the network is not affected in any way in this situation), and (ii) *permanent link failures* [3] defined as the definitive disappearance of an edge between two nodes of the graph. Results obtained indicate that irregular networked automata show an outstanding robustness and are more tolerant to faults compared to lattice CAs for the density and synchronization tasks [4].

3 Hill-Climbing Search for Network Structures

This chapter presents the development of a standard hill-climbing search method used to generate network topologies for the DCT in connection with the majority rule. A hill-climbing (HC) algorithm [10] is used with the following main steps:

- Generate an initial network structure (denoted by $\mathcal{C}$)
- While the maximum number of iterations has not been reached:
  (i) Apply mutation to $\mathcal{C}$ and obtain a new structure $\mathcal{N}$
  (ii) If $\mathcal{N}$ has a better fitness than $\mathcal{C}$ for DCT, then replace $\mathcal{C}$ with $\mathcal{N}$
- Return $\mathcal{C}$ as the final network structure solution.

It should be noted that this model does not exactly follow a random mutation HC (RMHC) approach because the mutation at step (i) is applied with a certain probability for each node in the network and not for only one randomly selected node (position) as in RMHC. Also, the HC model developed in this chapter differs from a steepest-ascent HC (SAHC) algorithm where mutation is systematically applied to each node and the best evaluated individual is selected each iteration.

The developed HC method uses a single individual which is a network encoded as an array of integers representing nodes and a list of links for each node. The initialization process starts from regular rings normally used to construct small-world networks [17, 19]. In the proposed approach, the initial network $\mathcal{C}$ is obtained

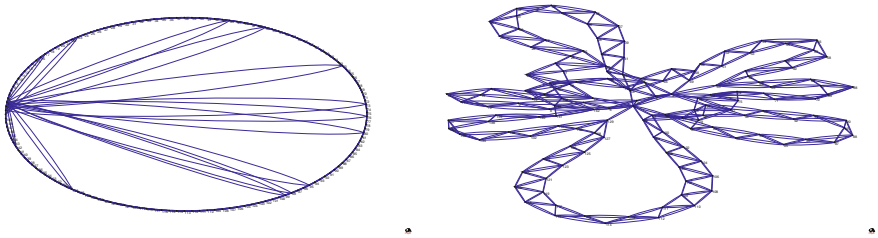


Fig. 1 Example of initial network used as a start point in the HC method. The network is drawn by Pajek software [22] using a circular layout as well as a representation by separate components

in the same way described in [17]. Each node is connected to its 4 closest neighbors in the ring and an edge is added with a probability of 0.1 for each node.

The fitness function for a network structure takes into account 100 initial configurations with uniformly distributed densities. The CA of size $N = 149$ is iterated using the majority rule (described in the previous section) based on the current network structure which gives the neighborhood of each cell. The maximum number of time steps is set to $M = 2N$. The fitness measures the fraction of initial configurations for which the majority rule in connection with the network topology leads to a correct classification. This fitness function is computationally less expensive compared to the performance of a rule which is measured over 10^4 initial configurations selected from an unbiased distribution. It should be noted that all these parameters (e.g. size of CA, time steps, number of initial configurations) are set to the same values used in most DCT studies [11, 13, 17].

Mutation is applied for each node of the network with a probability of 0.5 as explained in [17]. An edge from the selected target node to another randomly selected destination node is either added or removed with probability 0.5.

For the experiments presented in the current chapter, the maximum number of iterations is set to 100. This means that a network $\mathcal{C}$ is generated by the HC method using 100 successively applied mutations to the current best evaluated individual. The number of HC runs considered is 10. The best performance obtained is 0.8153 and the lowest performance is 0.7768. In three runs, the best performance was around 0.79 while in all other runs the value of the performance was above 0.8.

Figure 1 presents one of the initial networks generated in the first step of the HC algorithm, while Fig. 2 shows the structure of the best network generated after 100 HC iterations. Both networks have been drawn using the freely available Pajek software [22]. The performance of the initial network (Fig. 1) is 0.3281 and the performance of the final network (Fig. 2) is 0.8153.

An interesting behavior observed in the HC runs is that the algorithm usually starts with an initial configuration with a relatively high fitness of around 0.8 and low performance below 0.5. This fitness is improved on average during the first 10–15 iterations to the maximum value of 1 with a corresponding performance normally above 0.75. This behavior suggests the need for a better sustained search process potentially through the use of several individuals evolved in parallel and a

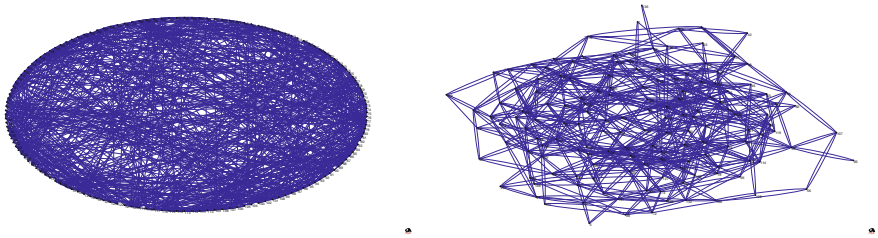


Fig. 2 Best network generated using the HC method. The performance for DCT is 0.8153. The network is drawn by Pajek software [22] using a circular layout and a representation by separate components

diversification procedure. It is expected that the use of a more complex search process can produce even more performant network based CAs. Nevertheless, the aim of this study is to emphasize the quality of networks generated by a simple HC algorithm for the DCT. The performance of the best networks obtained is further investigated in the next section under a dynamic framework.

4 Analysis of Network Dynamics

In order to evaluate the robustness of the obtained network topologies, we investigate their performance for DCT under dynamic changes understood as random removal or addition of network links. It should be noticed that we also consider the possibility of adding an edge between two nodes as opposed to permanent link failures tested in [3] where only edge removals are considered.

The number of steps with some changes performed in the network equals the number of edges in the network. In the current dynamics analysis, we tested the robustness of the best network generated by the HC method in 10 runs (see Fig. 2). This network has 354 edges and a performance of 0.8153 for DCT. Each network dynamics test starts with this network and each step performs a change in the network and evaluates the performance of the modified network for DCT. This process leads to the construction of 354 networks with their performances which will be analysed in this section.

The following three types of changes are considered at each step:

- *Case 1*: Each step, a randomly selected link is removed with the fixed probability p_r. If a link removal does not take place, a new random link is added with probability $p\_a = 1 - p\_r$. This means that at each step, either a link removal or an addition takes place. The considered values for p_r range from 0.1 to 0.9.
- *Case 2*: Each step, a randomly selected link is removed with the fixed probability p_r and a link is added with probability $p\_a = 1 - p\_r$. This means that at each step, a link removal, a link addition or both will take place. The considered values for p_r range from 0.1 to 0.9.

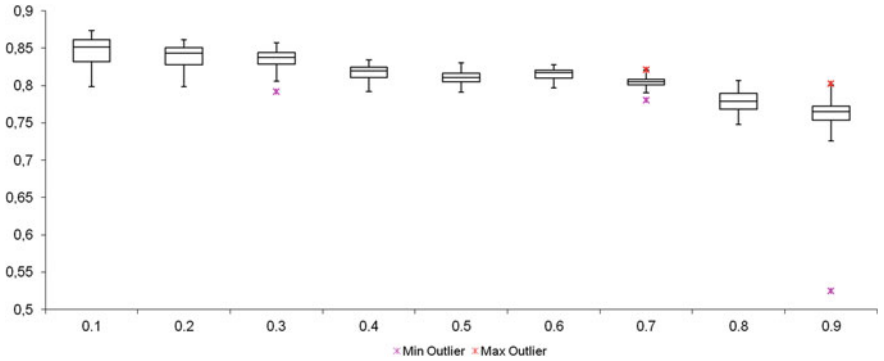


Fig. 3 Boxplot representation of results for case 1

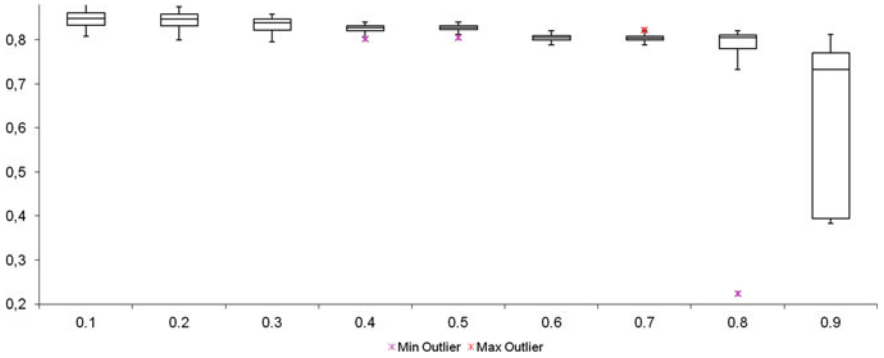


Fig. 4 Boxplot representation of results for case 2

- *Case 3*: This is the most general dynamic test in which the probabilities of removing and adding an edge are independent. Each step, a randomly selected link is removed with probability p_r and a new link is added with probability p_a. The considered values for these probabilities are 0.2, 0.4, 0.6 and 0.8, which leads to 16 (p_r, p_a) probability pairs analysed.

Figures 3 and 4 present the boxplot results for the first two cases considered.

A more stable performance is obtained when p_r is centered around 0.5 whereas the presence of outliers and more fluctuations on performance are observed when p_r and p_a are towards the extreme values allowed. Nevertheless, the performance of perturbed networks for the density task remains certain with most values above 0.8. This result is further emphasized in Figs. 5 and 6 which depict the maximum and minimum performance obtained for all considered probability values. The horizontal axis in these figures represents the value of p_a in descending order and the vertical axis is the CA performance. The value of p_r can be automatically deducted as it is always $1 - p\_a$ and not independent from it (as in dynamics test case 3). For case 1 (see Fig. 5), the maximum performance is above 0.85 and is obtained when

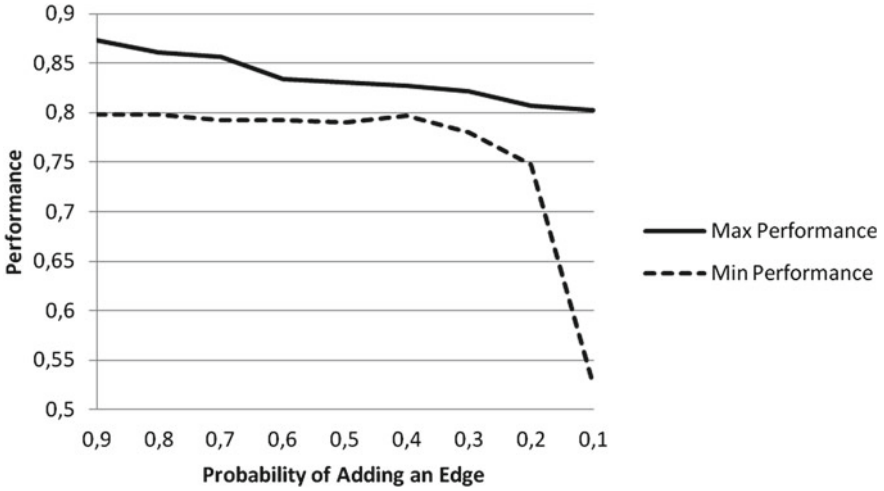


Fig. 5 Maximum and minimum CA performance obtained for networks perturbed according to dynamics test case 1. The x axis gives the p_a value and the corresponding p_r is $1 - p\_a$

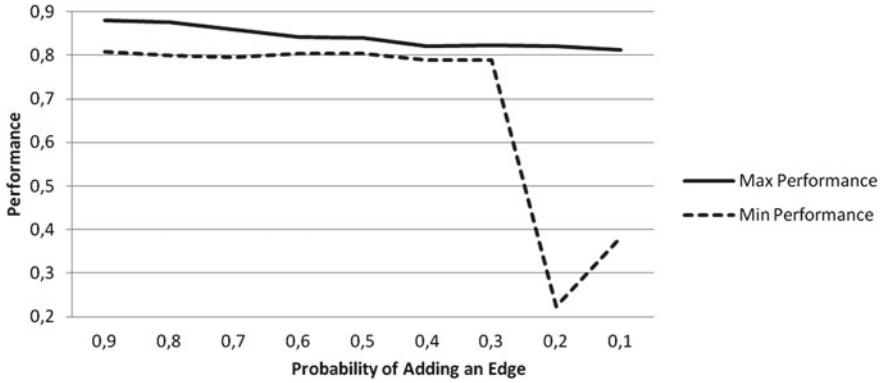


Fig. 6 Maximum and minimum CA performance obtained for networks perturbed according to dynamics test case 2. The x axis gives the p_a value and the corresponding p_r is $1 - p\_a$

$p\_a = 0.9$. The minimum performance stays around 0.8 for p_a value above 0.4 and then drops with the p_a. For case 2, the performance obtained has maximum values above 0.8 and minimum values around or slightly below 0.8 for p_a values of 0.3 or above. The robustness of the network is even more stable compared to the case 1 as now both edge additions and removals are possible at each step.

Figure 7 depicts the average performance of networks perturbed according to dynamics test case 3. In this case, the p_a and p_r values are independent so the x axis in Fig. 7 represents the p_a value while each line is for a different p_r value. The vertical axis represents the average performance obtained. The worst performance of around 0.78 is obtained when $p\_a = 0.4$ and $p\_r = 0.8$. Nevertheless, even this

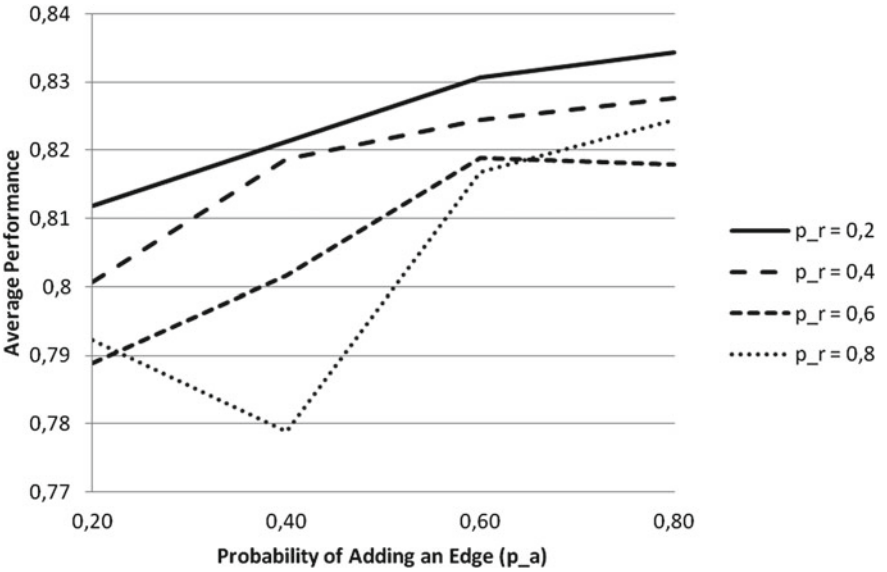


Fig. 7 Average performance obtained for independent p_a and p_r values in dynamics test case 3

lowest obtained performance is able to trigger a good classification result and it can not be considered as affecting the robustness of the network. As a trend, when the probability of removing an edge is high and the probability of adding one is low, the influence on performance is higher. Moreover, when the performance of adding an edge is high and the probability of removing one is low, we obtain networks with high average performances above 0.83, which is a better value even compared to the initial network. In fact, the maximum performance value obtained was 0.86 when $p\_a = 0.8$ and $p\_r = 0.2$.

The network dynamics analysis clearly supports a good robustness of the obtained networks as CA topologies in the DCT. Moreover, the results suggest the potential to obtain more performant networks by further search as the performance was improved from 0.8153 for the initial network to values as high as 0.86 for some perturbed versions of the starting point network.

5 Conclusions

A hill-climbing approach to generate network topologies for CAs in connection with DCT has been presented and analysed. It has been shown that a straightforward HC search algorithm is able to trigger good performances for network based CAs. Furthermore, dynamics of the obtained networks have been analysed and experiments

confirm the robustness of the networks which are able to maintain and even improve the performance of CA for DCT.

Future work focuses on using and extending as necessary the HC method to be used in connection with other search methods in order to facilitate the exploration of the search space. Mixed CA rules and topologies will also be investigated for density classification as well as other computational CA tasks.

Acknowledgments This research is supported by Grant PN II TE 320, Emergence, auto-organization and evolution: New computational models in the study of complex systems, funded by CNCS Romania.

References

1. Andre, D., Bennett III, F.H., Koza, J.R.: Discovery by genetic programming of a cellular automata rule that is better than any known rule for the majority classification problem. Proceedings of the First Annual Conference on Genetic Programming. GECCO '96, pp. 3–11. MA, USA, MIT Press, Cambridge (1996)
2. Baetens, J.M., De Baets, B.: Topology-induced phase transitions in totalistic cellular automata. Physica D **249**, 16–24 (2013)
3. Darabos, C., Giacobini, M., Tomassini, M.: Performance and robustness of cellular automata computation on irregular networks. Adv. Complex Syst. **10**, 85–110 (2007)
4. Darabos, C., Tomassini, M., Di Cunto, F., Provero, P., Moore, J.H., Giacobini, M.: Toward robust network based complex systems: from evolutionary cellular automata to biological models. Intelligenza Artificiale **5**(1), 37–47 (2011)
5. Ferreira, C.: Gene Expression Programming: A New Adaptive Algorithm for Solving Problems. Complex Systems **13**(2), 87–129 (2001)
6. Gog, A., Chira, C.: Cellular automata rule detection using circular asynchronous evolutionary search, HAIS 2009. LNCS **5572**, 261–268 (2009)
7. Juille, H., Pollack, J.B.: Coevolutionary learning and the design of complex systems. Adv. Complex Syst. **2**(4), 371–394 (2000)
8. Land, M., Belew, R.K.: No perfect two-state cellular automata for density classification exists. Phys. Rev. Lett. **74**(25), 5148–5150 (1995)
9. Mitchell, M., Crutchfield, J.P., Das, R.: Evolving cellular automata with genetic algorithms: A review of recent work. In: Proceedings of the First International Conference on Evolutionary Computation and Its Applications (EvCA'96). Russian Academy of Sciences (1996)
10. Mitchell, M., Forrest, S.: Royal Road functions. In: Back, T., Fogel, D., Michalewicz, Z. (eds.) Handbook of Evolutionary Computation. Oxford University Press, Oxford (1998)
11. Mitchell, M., Thomure, M. D., Williams, N. L.: The role of space in the Success of Coevolutionary Learning. In: Proceedings of ALIFE X —The Tenth International Conference on the Simulation and Synthesis of Living Systems (2006)
12. de Oliveira, P.P.B., Bortot, J.C., Oliveira, G.: The best currently known class of dynamically equivalent cellular automata rules for density classification. Neurocomputing **70**(1–3), 35–43 (2006)
13. Oliveira, G.M.B., Martins, L.G.A., de Carvalho, L.B., Fynn, E.: Some investigations about synchronization and density classification tasks in one-dimensional and two-dimensional cellular automata rule spaces. Electron. Notes Theor. Comput. Sci. **252**, 121–142 (2009)
14. Packard, N.H.: Adaptation toward the edge of chaos. In: Shlesinger, M.F. (ed.) Dynamic Patterns in Complex Systems, World Scientific, Singapore pp. 293–301 (1988)
15. Pagie, L., Mitchell, M.: A comparison of evolutionary and coevolutionary search. Int. J. Comput. Intell. Appl. **2**(1), 53–69 (2002)

16. Tomassini, M., Venzi, M.: Evolution of Asynchronous Cellular Automata for the Density Task. Parallel Problem Solving from Nature—PPSN VII. Lecture Notes in Computer Science, Springer, Berlin / Heidelberg **2439**, 934–943 (2002)
17. Tomassini, M., Giacobini, M., Darabos, C.: Evolution and dynamics of small-world cellular automata. Complex Syst. **15**, 261–284 (2005)
18. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'smallworld' networks, Nature **393, 440–442 (1998)**
19. Watts, D.J.: Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton University Press, Princeton (1999)
20. Watts, D.J.: Six degrees: The Science of a Connected Age. Gardner's Books, New York (2003)
21. Wolfram, S.: Theory and Applications of Cellular Automata, Advanced Series on Complex Systems, World Scientific Publishing, Singapore, p. 9128 (1986).
22. Pajek Software, http://www.pajek.imfm.si/doku.php

Autocorrelated Random Walks and Entropy

Rudolf Hanel and Stefan Thurner

Abstract We show with two simple examples, one—an autocorrelated random walk, the other—an accelerated random walk, that two processes that are fundamentally different on a microscopical level, so different in fact that the two processes implement different types of entropic concepts, still can be indistinguishable from a probabilistic point of view, i.e. all finite moments of the two processes may coincide. The immediate consequence of this observation is that entropy primarily is a property associated with the structure of phase-space rather than a consequence of specific observable distribution functions.

Keywords Non-Markovian · Non-ergodic · Path-dependent processes · (c,d)-entropies

1 Introduction

In the statistical theory of thermodynamic systems the concept of Entropy plays a fundamental role. In the work of Boltzmann the entropy measures the number of states, W, a system can be found in on a logarithmic scale, i.e.

$$S \propto \log(W)\,. \tag{1}$$

R. Hanel · S. Thurner (✉)
Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria
e-mail: stefan.thurner@meduniwien.ac.at

S. Thurner
Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

R. Hanel
e-mail: rudolf.hanel@meduniwien.ac.at

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_28,

For a wide class of N-particle systems the entropy $S(N) \sim NS(1)$, i.e. W essentially grows exponentially with the system size in terms of particles, which is to say the entropy is *extensive*. This exponential behavior is ubiquitous for a wide class of systems and basically follows from the multinomial structure of sampling a system of N identical processes, i.e. from the total symmetry of the N-particle distribution function (exchanging two identical particles does not change the distribution function), i.e. the invariance under the permutation group. If ϕ are some intensive system variables and $W(N) \sim (\sum_\phi w(\phi))^N$ for some function $w(\phi) \geq 0$, then it becomes clear that in the limit $N \to \infty$ the function $W(N)$ essentially behaves like $w(\hat{\phi})$, where w is maximal for the argument $\hat{\phi}$. In a binomial process, e.g. a spin system, this means that if the system is constrained in such a way that the probability of finding a spin to be up is p, then the number of states with pN spins up, using Stirling's approximation is essentially given by $\binom{N}{Np} \sim 1/(p(1-p))^N$ and $S = \log \binom{N}{Np} \propto N$.

This is the key to deriving the functional form of the entropy for the associated maximum entropy principle (MEP) [1–3], which for the probabilities p_i to observe state $i = 1, \ldots, W$ in an extensive systems takes the mathematical form

$$S[p] = -k_B \sum_{i=1}^{W} p_i \log(p_i) \tag{2}$$

where k_B (which we will set equal to one) is the Boltzmann constant. In this functional form we will refer to the entropy as Shannon's entropy. While Boltzmann's entropy is the same for non-extensive systems, i.e. $\log(W)$, only that it is no longer proportional to N, the form of the maximum entropy functional $S[p]$ is only of Shannon's form for extensive systems. In non-extensive systems and some *asymptotically* extensive systems this situation changes [4–8]. Boltzmann's entropy relies on an universal idea - the properties of a large systems ($N \gg 1$) are dominated by the properties of the most abundant type of states the system can be found in and therefore Boltzmann's entropy is always the same expression, whether a system is extensive or not: $S \propto \log(W)$ - In the non-extensive case the proportionality factor can become some function $f(N)$ that differs from N. In the simplest cases this can be a power $f(N) = N^\alpha$. We will study such a system below.

From this observation distinct notions of generalized entropies can be derived, which are related but not necessarily equivalent. One possibility is to find an entropy functional $S_\Lambda[p]$, that generalizes Boltzmann's approach to $S_\Lambda \propto \Lambda(W)$, for Λ being some other monotonically increasing function, called the *generalized logarithm*, in such a way that $\Lambda(W) \propto N$, i.e. the new entropy notion is again extensive. This possibility has already been pointed out by Tsallis (e.g. [9, 10]). In [4] it has been shown that in fact $W(N)$ determines the form of the functional $S_\Lambda[p]$. Moreover, it has been shown how entropic functionals S_Λ can be classified in terms of their asymptotic behavior [5]. This work uses the fact that Shannon's entropy is uniquely determined by 4 axioms (Shannon-Khinchin axioms SK1-SK4, [11, 12]) and that non-extensive systems violate one axiom often referred to as the *composition axiom*. The *extensive* generalized entropy functional $S_\Lambda[p]$ however, usually is

not the appropriate MEP functional associated with the system and the authors are currently investigating the question how generalized MEP can be derived for various classes of non-extensive, aging systems.

While these general results will be published elsewhere, we will focus on a particular simple example of a non-extensive process that exemplifies two things: (i) strong long-range correlations, breaking the total symmetry of the sampling process, can cause phase-space to grow non-exponentially and therefore generalized entropies become necessary and (ii) the probabilistic description of such a system may be equivalent to the probabilistic description of another system, which is in fact extensive (i.e. the distribution function does not tell us whether a system is extensive or not).

In the following we will look at a class of autocorrelated random walks which require generalized entropies. We show that these random walks can be used to construct autocorrelated versions of the Wiener measure. In the continuum limit these autocorrelated measures are probabilistically equivalent to Fokker-Planck diffusion processes with time dependent diffusion constant.

2 A Class of Autocorrelated Random Walks and its Continuum Limit

Let us begin with defining a discrete process

$$\Delta W_{\Delta t}(n\Delta t) = \omega_n \Delta x \,, \tag{3}$$

where $\omega_n \in \{-1, +1\}$ and the expectation value of $\langle \omega_n \rangle = 0$ for all n. The construction closely follows the construction of the Wiener process. However, instead of having $\langle \omega_n \omega_m \rangle = \delta_{nm}$, where δ_{mn} is Kronecker's delta function, ω_n are correlated in the following way

$$\langle \omega_n \omega_m \rangle = 1 \text{ if } \exists k\colon k < n^\alpha, m^\alpha < k + 1 \tag{4}$$

and $\langle \omega_n \omega_m \rangle = 0$ otherwise. Clearly one requires $0 < \alpha \leq 1$. This means that with passing time the process will keep walking persistently in one direction for increasing numbers of time steps. The process can not freely decide whether to go left or right at any time step.

One process for instance that asymptotically has $\alpha = 1/2$ can decide in the beginning whether to go left or right. After making this decision it goes one step, say to the right. Then it faces another decision. Now having made this decision, say left, the walker goes in this direction, but not one but two steps. With any new decision the walker has to go first three steps in the same direction, then four, fife, and so on. For an uncorrelated random walk, the number of different paths a walker can take is given by $\log W(t)/\log 2 = N(t)$, where $N(t) = [t/\Delta t]$. The function $[x]$ denotes the next closest integer smaller than x. In contrast, for the correlated random walk

the number of possible configurations after time t is given by $\log W(t)/\log 2 = [(t/\Delta t)^\alpha] \sim N(t)^\alpha$. So clearly the correlated random walk is non-extensive and requires a generalized extensive entropy [4].

In the next step we can construct the continuum limit of the autocorrelated random walk to obtain the autocorrelated analogon for the Wiener process in a Riemann-Stieltjes sense. For this we look at the second moment $\langle W_{\Delta t}(t)\rangle$ and, with χ being the characteristic function, obtain

$$\begin{aligned}
\langle W_{\Delta t}(t)^2\rangle &= \sum_{m=1}^{[t/\Delta t]}\sum_{n=1}^{[t/\Delta t]}\langle\omega_m\omega_n\rangle\Delta x^2 \\
&= \sum_{k=0}^{\infty}\left(\sum_{m=1}^{[t/\Delta t]}\sum_{n=1}^{[t/\Delta t]}\chi(k\le m^\alpha, n^\alpha < k+1)\right)\Delta x^2 \\
&= \sum_{k=0}^{\infty}\left(\sum_{n=1}^{[t/\Delta t]}\chi(k\le n^\alpha < k+1)\right)^2\Delta x^2 \\
&= \sum_{k=0}^{\infty}\theta\left(\left[(t/\Delta t)^\alpha\right]-k\right)\left(k^{1/\alpha}\left((1+1/k)^{1/\alpha}-1\right)\right)^2\Delta x^2 \\
&\sim \sum_{k=0}^{[(t/\Delta t)^\alpha]}\frac{1}{\alpha^2}k^{2/\alpha-2}\Delta x^2 \\
&\sim \frac{1}{\alpha(2-\alpha)}\Delta x^2\, k^{2/\alpha-1}\Big|_0^{(t/\Delta t)^\alpha} \\
&\sim \frac{1}{\alpha(2-\alpha)}\Delta x^2\left(\frac{t}{\Delta t}\right)^{2-\alpha} \\
&\stackrel{!}{=} t^\beta\,, \qquad (5)
\end{aligned}$$

where in the last line we impose that that t^β is the result. This allows us to define the continuum process $W_\beta(t) = \lim_{\Delta t\to 0} W_{\Delta t}(t)$ as the Riemann-Stieltjes limit of the above process by identifying $\beta = 2-\alpha$ (this means $1 \le \beta < 2$) and by defining $\Delta x = \Delta x_\beta(\Delta t)$ as

$$\Delta x_\beta(\Delta t) = \sqrt{\beta(2-\beta)}\Delta t^{\frac{\beta}{2}}\,, \qquad (6)$$

which then implies $\langle W_\beta(t)^2\rangle = t^\beta$. Clearly $\beta = 1$, i.e. $W_1(t)$, corresponds to the usual Wiener-process with the well known relation $\Delta x^2 \sim \Delta t$. Autocorrelations, as we have just shown, can fundamentally change this scaling relation.

In the next step we can look at all finite moments $\langle W^n(t)\rangle$ and find an astonishing equivalence. The autocorrelated process is probabilistically equivalent to a diffusion process $dX(t) = B(t)dW_1(t)$, with time dependent diffusion constant $D = B^2$.

3 Finite Moments and Equivalence

It is clear from $\langle\omega_n\rangle = 0$ that all odd moments of the process $W_\beta(t)$ are vanishing, $\langle W_\beta(t)^n\rangle = 0$ for $n = 1, 3, 5, \ldots$. For even moments these moments can be computed

and one finds

$$\begin{aligned}\langle W_{\Delta t}(t)^{2n}\rangle &= \langle \int_0^t dW_\beta(t_1)\int_0^t dW_\beta(t_2)\cdots\int_0^t dW_\beta(t_{2n})\rangle \\ &= \sum_{m_1,m_2,\ldots,m_{2n}=0}^{[t/\Delta t]} \langle \omega_{m_1}\omega_{m_1}\cdots\omega_{m_{2n}}\rangle \Delta x^{2n} \\ &= \frac{(2n)!}{n!2^n}\langle W_\beta(t)^2\rangle^n \\ &= \frac{(2n)!}{n!2^n}t^{\beta n} \end{aligned} \tag{7}$$

Let us compare these results with the ones obtained for an accelerated process

$$dX(t) = t^\gamma dW_1(t) \tag{8}$$

and get for the second moment:

$$\begin{aligned}\langle X(t)^2\rangle &= \langle \int_0^t dW_1(t_1)\int_0^t dW_1(t_2)t_1^\gamma t_2\gamma\rangle \\ &= \int_0^t dt_1 t_1^{2\gamma} \\ &= \frac{t^{2\gamma+1}}{2\gamma+1} \end{aligned} \tag{9}$$

This means that $2\gamma + 1 = \beta$, i.e. $\gamma = (\beta - 1)/2$ and, identifying the accelerated process to be

$$d\bar{W}_\beta(t) \equiv \underbrace{\sqrt{\beta}t^{\frac{\beta-1}{2}}}_{\equiv B(t)} dW_1(t)\,, \tag{10}$$

one sees that $\bar{W}_\beta(t)$ possesses exactly the same finite moments as $W_\beta(t)$, even though both processes are fundamentally different in terms of their microscopic structure.

4 Discussion

Evidently both the autocorrelated and the accelerated process are identical in a weak probabilistic sense; i.e. all moments of the two processes are identical. Yet, the two processes differ completely in their microscopic behavior. The number of possible configurations, W, grows exponentially for the accelerated process, which therefore this is extensive, and subject to Shannon's entropy. In case of the autocorrelated process W grows sub-exponentially. The process is non-extensive and requires generalized notions of entropy.

The most striking differences of the two processes are maybe the following: (i) while in the autocorrelated case the random walker always walks at constant speed v_β which scales like

$$v_\beta = \frac{\Delta x}{\Delta t} = \sqrt{\beta(2-\beta)}\Delta t^{\frac{\beta}{2}-1}, \tag{11}$$

the velocity of the accelerating walker $\bar{v}_\beta$ increases with time t and scales with Δt in the following way

$$v_\beta(t) = \frac{\Delta x}{\Delta t} = \sqrt{\frac{\beta}{\Delta t}} t^{\frac{\beta-1}{2}}. \tag{12}$$

The second striking feature is that, (ii) while the accelerated process chooses a new direction with equal probability at every time step, the autocorrelated process becomes more and more persistent with time, i.e. the probability of the autocorrelated process to change direction in fact converges to zero in the continuum limit.

In other words, if we consider the random variables of the two processes not to be the choices $\omega_n = \pm 1$ to either go left or right, but the choices η_n, whether the walker changes direction after the n'th step or not, then we can write $\omega_{n+1} = \eta_n \omega_n$. If $\eta_n = 1$ the walker continues in the direction it is heading at the moment. If $\eta = -1$ the walker reverses its direction. We can now examine the probability $p(k) = P(\sum_{n=1}^{N} \chi(\eta_n = -1) = k)$ for the two processes and since in the accelerated case all ω_n are independent and $\langle \omega_n \rangle = 0$ it follows that also $\langle \eta_n \rangle = 0$ and therefor

$$p_{\text{accelerated}}(k) = \binom{N}{k} p^k (1-p)^{N-k}, \tag{13}$$

is binomially distributed and the probability of direction-reversal is $p = 1/2$ for all N. In the autocorrelated case however, $p(k)$ ceases to be binomially distributed and the probability of direction-reversal $p \to 0$ as $N \to \infty$. This on the other hand implies that the probability of sequences of N decisions on changing direction that contain k decisions to change direction and $N - k$ not to, are not invariant under permutation of the occurrences of those events. The hallmark of systems that require generalized entropies is exactly this broken invariance. In fact this is the starting-point of an analysis that allows to identify classes of systems where this invariance can be considered as deformed rather than completely broken. Such systems are characterized by the fact that every new realization of the process will break symmetry and become more and more persistent, similarly to the autocorrelated random walk.

5 Conclusions

We have shown with the simple examples of autocorrelated and an accelerated random walks that although both processes are fundamentally different on a microscopical level, so different in fact that the two processes have to be described with

different types of entropies, that they are indistinguishable with respect to all their finite moments. The main conclusion that we have to draw from this is the following. The weak equivalence of different processes implied by the identity of all moments does not allow to infer that two processes have to be identical with respect to the entropy of the two systems! The entropy concept is therefore related to the microstructure of a process and the internal constraints that break the total symmetry of the joint probability distribution function describing the system.

References

1. Jaynes, E.T.: Probability Theory: The Logic of Science, Cambridge University Press, pp. 351–355 (2003)
2. Jaynes, E.T.: Phys. Rev. Ser. **II**(106), 620–630 (1957)
3. Jaynes, E.T.: Probability Theory: The logic of the science. Cambridge University Press, Cambrigde UK (2003)
4. Hanel, R., Thurner, S.: Euro. Phys. Lett. **96**, 50003 (2011)
5. Hanel, R., Thurner, S.: Euro. Phys. Lett. **93**, 20006 (2011)
6. Hanel, R., Thurner, S.: Phys. A **380**, 109–114 (2007)
7. Thurner, S., Hanel, R.: In complexity, metastability and nonextensivity. AIP **965**, 68–75 (2007)
8. Hanel, R., Thurner, S., Gell-Mann, M.: PNAS **108**, 6390–6394 (2011)
9. Tsallis, C.: Introduction to Nonextensive Statistical Mechanics—Approaching a Complex World. Springer, New York (2009)
10. Tsallis, C., Gell-Mann, M., Sato, Y.: Asymptotically scale-invariant occupancy of phase space makes the entropy Sq extensive Proceedings of the National Academy of Sciences, vol. 102, pp. 15377–15382 (2005)
11. Khinchin, A.I.: Mathematical foundations of information theory. Dover Publ, New York (1957)
12. Shannon, C.E.: Bell Syst. Tech. J. **27**(379–423), 623–656 (1948)

Complex Network Construction Based on SOMA: Vertices In-Degree Reliance on Fitness Value Evolution

Lenka Skanderova, Ivan Zelinka and Petr Saloun

Abstract This paper deals with complex network constructing with using evolutionary algorithm SOMA AllToOne version. The main goal is to visualize complex networks developing and analyse their properties, especially in-degrees and their realiances on fitness value evolution. Thank this analysis we can make an analysis of the populations evolutions during the evolutionary algorithm.

Keywords Complex networks · Evolutionary algorithm · SOMA · AllToOne · Vertex in-degree

1 Introduction

Evolutionary algorithms are based on three basic principles—natural selection, crossing and mutation. The population of individuals plays the main role in these algorithms. Each individual has its own parameters and fitness value. Each parameter has its lower and upper bound. It is not possible to cross these bounds. Fitness value says how this individual is good in the population. Evolution algorithms are used for optimization, so the global minimum or maximum of the cost function (fitness function) is ussually searched. New population is generated randomly at the beginning of the algorithm. Word "randomly" means the individuals parameters are generated randomly in their lower and upper bounds. The fitness of each individual is computed.

L. Skanderova (✉) · I. Zelinka · P. Saloun
Department of Computer Science, VSB -Technical University of Ostrava, 17. listopadu/2172, 708–33 Ostrava - Poruba, Czech Republic
e-mail: lenka.skanderova@vsb.cz

I. Zelinka
e-mail: ivan.zelinka@vsb.cz

P. Saloun
e-mail: petr.saloun@vsb.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_29,

Then the evolution process can begin. The individuals are crossed and mutated. Best survive while worse die [1]. In this paper, SOMA AllToOne version has been chosen from the family of the evolutionary algorithms. In connection with complex networks evolutionary algorithms are mentioned for example in [2–5].

SOMA differs from typical evolutionary algorithms principle, for example from differential evolution (DE), by the main idea of the algorithm. In DE new offspring is created during the evolution. In SOMA the individuals migrate in the space of possible solutions. There is no offspring, individuals just change their positions. In this paper SOMA AllToOne has been used, because all individuals migrate to one individual—we call it *Leader*—and it is easy to observe the evolution of relationships between individuals and *Leaders* during complex network's construction. We can observe the "winners", individuals, which have been chosen as *Leaders* most often.

A network is a set of items, we call them vertices, whith connections between them, theese connections are called edges. We can mention for example World Wide Web, social networks, metabolic networks, food webs etc. It is known that the matematical graph theory is one of the fundamental pillars of discrete mathematic. Networks have also been studied in the social sciences, where typical social network studies address issue centrality and connectivity [6]. The vertices and edges can have many properties, which can be researched. There can be more types of vertices as well as edges in the network. Vertices can represent states, people or organizations. Edges relationships—e.g. friendship etc. Edges can have their weights and they can be directed or not directed. Graphs composed of directed edges are called digraphs. Graphs can contain hyperedges—edges, which join more than two vertices together [6]. In connection with genetic algorithm [7] has been written in 2013 [8] describes a bibliometric study of service innovation research based on complex network analysis. In [9] authors write about the study of emergence in small-group social networks [10] occupies by directed complex networks and in [11] new global synchronization analysis for complex networks is described.

2 Motivation

The main motivation of this paper is to observe reliance of the individuals in-degree values on fitness value evolution during the evolution process. The special attention has been devoted to the areas, where fitness value is stagnating.

3 Experiment Design

For experiments HP Pavilion dv7-6050 with processor Intel Core i7 with frequency 2 GHz, 4 GB RAM and graphic card AMD Radeon HD 6770M and Microsoft Visual Studio 2010 have been used. The experiments have beed processed by Mathematica 8, Gephi 0.8.2 and Gnuplot 4.6. As it was mentioned above, SOMA AllToOne has been

Table 1 SOMA AllToOne setting

Parameter	Value
NP	100
D	50
$Migrations$	300
PRT	0.1
$PathLength$	3
$Step$	0.11

used as the evolutionary algorithm. The global minimum of the function has been searched. And Schwefel's function has been chosen as a testing function. Schwfel's global minimum is $f(x) = 418.9829 \times D$, where D denotes dimension. For this function the set of complex networks has been constructed. Each migration of the algorithm meant one step in the complex network developing. Each step has been registered and analyzed. 300 graphs have been constructed. Exact setting of the algorithm is reffered in Table 1. In Table 1 NP means number of individuals in the population, D—dimension, $Migrations$ number of migration cycles, PRT means perturbation, perturbation vector is created according to this parameter, $PathLength$ says how far the actual individual will stop from the *Leader* and $Step$ means "grain" of individual's path. The construction of the graph for SOMA AllToOne has been implemented like this:

- One vertex in the graph meant one individual in evolutionary algorithm.
- In each generation the best individual has been found. Best individual means the individual with the best (minimal) fitness.
- The best individual has been supported by each other individual. So from the view of the graph it had the connection with each other vertex. In this paper oriented graph has been used for visualizations. If the individual is supported by others in the evolutionary algorithm, number of edges leading to the vertex, which visializes this individual will increase and the in-degree of this vertex will increase too. Thanks this mechanism we can observe the complex network evolution, which individual has become *Leader* the most often, if there is the phenomenon "rich become to be richer" etc.

4 Results

From the view of results we have been interrested in phenomenon "rich become to be richer" and especially a connection of vertex in-degree stagnation with evolution process stagnation in SOMA. In other words if in-degree of vertex will stagnate when *Leader* does not change.

We can observe the phenomenon "rich become to be richer" by the individual number 5. Individual number 5 has the highest in-degree. Its value is 5148. Next

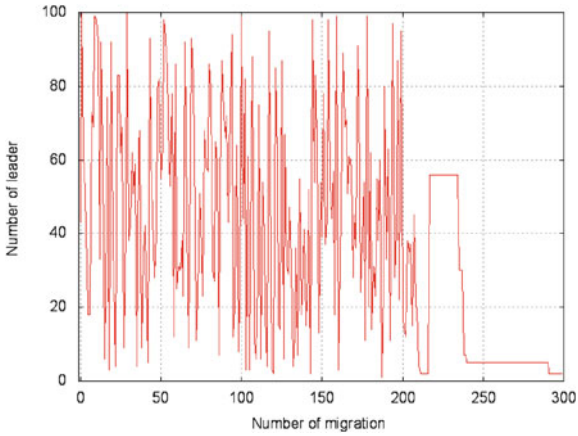


Fig. 1 Schwefel's function, *Leaders* visualization during the evolution process. It cen be observed that individuals number 2, 5 and 56 are the richest

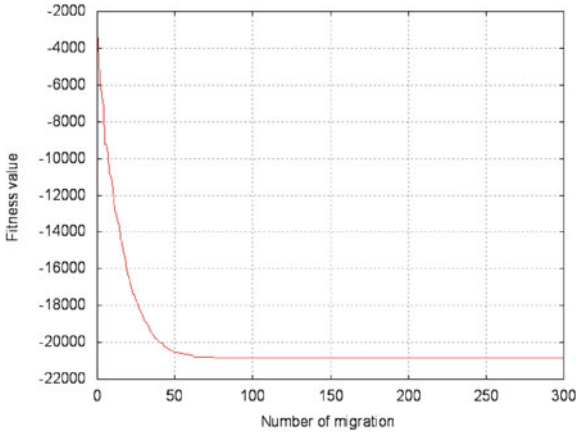


Fig. 2 Schwefel's function fitness evolution during evolution process

rich individuals are individual number 2 with in-degree 1683 and 56 with in-degree 1782. Individual number 2 is the last *Leader* in the evolution process. In Fig. 2 the fitness value evolution is depicted. The fitness stagnates from the migration number 100 with the value −20830.7. From the Fig. 1 it is clear that *Leaders* have changed many times during the evolution after 100th migration. From the migration number 212 *Leaders* change much slower than before.

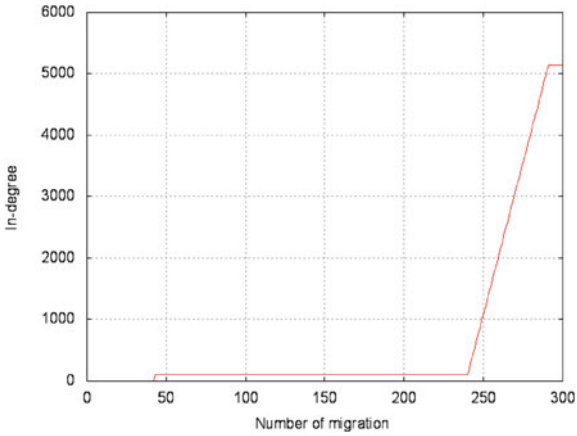


Fig. 3 Schwefel's function, individual number 5 in-degree visualization. We can observe in which migration cycle it has become *Leader* at first and its in-degree evolution

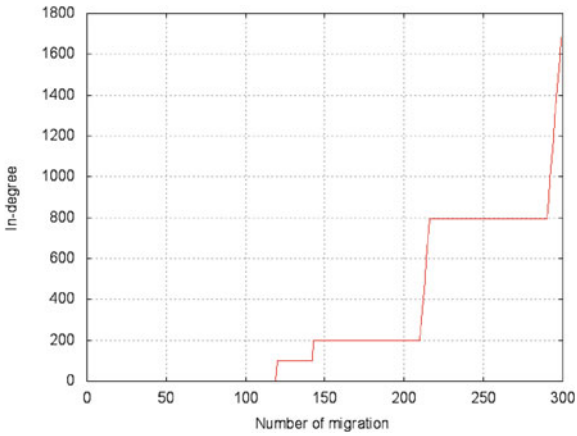


Fig. 4 Schwefel's function, individual number 2 in-degree visualization. We can observe in which migration cycle it has become *Leader* at first and its in-degree evolution

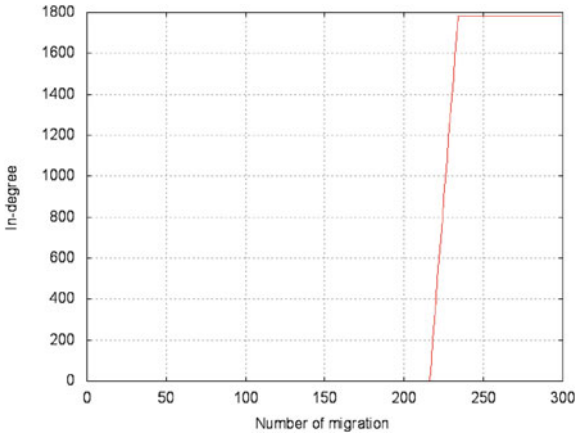


Fig. 5 Schwefel's function, individual number 56 in-degree visualization. We can observe in which migration cycle it has become *Leader* at first and its in-degree evolution

5 Conclusion

From the Sect. 4 we can make some conclustions:

- For Schwefel's function 300 migrations have sufficed to observe the phenomenon "rich become to be richer", in this case individual number 5 has become winner. On the other hand 300 migrations have not sufficed to reach the global minimum of the Schwefel's function which is mentioned in Sect. 3. This is probably caused by high dimension. However this count of migrations sufficed to observe vertices in-degree evolution connection with evolution stagnation in SOMA.
- The phenomenon "rich become to be richer" appeared by the individuals number 2 and 56 but their in-degree were much lower than in-degree of individual number 5.
- The Schwefel's function fitness stagnates from the migration number 100, its value is −20830.7. When we round the fitness value to integer, the fitness will stagnate from the migration number 92 with the value −20831, see Fig. 2. If we look at Fig. 1 we will see that from the migration number 100 till 208 the phenomenom "rich become to be richer" can not be observed and *Leaders* change very quickly too. From the migration number 212 *Leaders* change much slower than before.
- In Figs. 3, 4 and 5 in-degree evolutions of the individuals numbers 5, 2 and 56 are depicted. Their in-degrees have been increasing very slowly till the last third of the evolution process. Then their in-degrees increased much faster. From these three individuals, individual number 5 has been chosen as the *Leader* first. In Fig. 3 we can see long in-degree stagnation around the migration cycle number 100, this situation lasts 198 migration cycles. The second richest individual is individual number 2. We can see in Fig. 4 there are three long in-degree stagnations during the evolution process.
- Individual number 56 is the third richest individual, it has become *Leader* 17 times. There is a big differences between individuals number 2 and 56, and individual number 5. Individual number 2 has become *Leader* 18 times, individual number 5617 times while individual number 552 times.
- *Leaders* change very quickly till the last third of the migration cycles. The individual number 2 has been chosen as the last *Leader*. As it is mentioned above the algorithm has not reached the global minimum of the Schwefel's function. It is probable that in-degrees will still increase and *Leaders* will change, but this process will be very slow.

In the future we would like to make next experiments with greater collections of individuals, other optimization functions and other evolutionary algorithms, e.g. DE, to investigate the complex network structure creation during the evolutionary algorithms.

Acknowledgments The following two grants are acknowledged for the financial support provided for this research: Grant Agency of the Czech Republic—GACR P103/13/08195S, by the Development of human resources in research and development of latest soft computing methods and their application in practice project, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational

Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic

References

1. Stoean, R., Stoean, C.: Modeling medical decision making by support vector machines, explaining by rules of evolutionary algorithms with feature selection. Expert Syst. Appl. **40**, 2677–2686 (2013)
2. Li, Y.F., Sansavini, G., Zio, E.: Non-dominated sorting binary differential evolution for the multi-objective optimization of cascading failures protection in complex networks. Reliab. Eng. Syst. Safe. **111**, 195–205 (2013)
3. Klimkova, E., Senkerik, R., Zelnka, I.: Visualization of giant connected component in directed network. MENDEL 2011–17th International Conference On Soft Computing, Mendel, pp. 486–491 (2011)
4. Tomsu, L., Zelinka, I.: Complex networks and evolutionary algorithms. MENDELL 2009, Mendel, pp. 55–61 (2009)
5. Zelinka, I. et al.: Evolutionary dynamics and complex networks. MENDEL 2012–18th International Conference On Soft Computing, Mendel, pp. 88–93 (2012)
6. Newman, M.E.J.: The structure and function of complex networks. Soc. Ind. Appl. Math. **45**, 167–256 (2003)
7. Liu, D.Y. et al.: Genetic algorithm with a local search strategy for discovering communities in complex networks. Int. J. Comput. Intell. Syst. **6**, 354–369 (2013)
8. Zhu, W.J., Guan, J.C.: A bibliometric study of service innovation research: based on complex network analysis. Scientometrics **94**, 1195–1216 (2013)
9. Lewis, T.G.: Cognitive stigmergy: a study of emergence in small-group social networks. Cogn. Syst. Res. **21**, 7–21 (2013)
10. Ma, Q., Lu, J.W.: Cluster synchronization for directed complex dynamical networks via pinning control. Neurocomputing **101**, 354–360 (2013)
11. Gong, D.W. et al.: New global synchronization analysis for complex networks with coupling delay based on a useful inequality. Neural Comput. Appl. **22**, 205–210 (2013)

Sentiment Analysis in Complex Adaptive Systems

Petr Šaloun, Ivan Zelinka and Martin Hruzik

Abstract The aim of this work is to present a new algorithm for the evaluation of sentiment in Czech language texts. The algorithm is based on a new dictionary and uses n-gram searching. For the creation of the dictionary, it was important to use language specific phrases and exceptions, which can completely change the final evaluation of a sentiment. The solution also includes automatic search for a new subjects (aspects) of evaluation and also searching for new words determining sentiment. A similar algorithm can also be applied to other languages. The work emphasizes the transformation of the acquired data into valuable information. Our experiment is realized in the experimental adaptive web system in e-learning content domain and in eShop domain. The success and benefits of the algorithm are also discussed in this text.

Keywords Sentiment analysis · Opinion mining · Lemmatization · Aspects mining · Wordnet

1 Introduction

The Internet is constantly replenished with new users who generate more and more content. The content conceals information of high value, and many companies, scientists, prefessionals and experts are trying to figure out how to obtain and use

P. Šaloun (✉) · I. Zelinka · M. Hruzik
VSB-Technical University of Ostrava, 17. listopadu 15, 70833 Ostrava, Czech Republic
e-mail: petr.saloun@vsb.cz

I. Zelinka
e-mail: ivan.zelinka@vsb.cz

M. Hruzik
e-mail: me@martinhruzik.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_30,

this valuable information. Subjective texts, which are important for their sentiment, are a part of this.

Sentiment analysis (opinion mining) is currently one of the most discussed topics. In the past, there have been several startups and a number of methods created with the same goal—get the sentiment from text. These methods are used for example in financial markets, where sentiment analysis helps with stock trading.

The most common application of sentiment analysis is in the area of consumer reviews of products and services. Opinion mining helps producers to find out what the customers think about their products and determine what they like and what they complain about. This information can help manufacturers to further develop a product, and subsequently, increase sales. Customers on the other hand, can see what other people refer to as an advantage or disadvantage and then decide whether to buy the product or not. There are numerous news items, articles, blogs, tweets etc. which are analyzed [1].

It is possible to find many different ways for doing sentiment analysis of English texts. Some of these methods have been already implemented for Chinese and Spanish texts. The other languages usually use the common solution—first, the text is converted into English, and then the analysis of sentiment is done. Each language has its own peculiarities, so the method that involves a translation is not always successful.

This text is focused on the procedure of sentiment analysis in adaptive systems based on Slavic languages and Czech language primarily. We have used my experience from e-shops and native knowledge of the Czech language to design the algorithm. In addition to analyzing data about specific products, the algorithm was also succesfully applied in the domain of e-learning, where we gathered students' feedback. Adaptive web systems in education domain are described in [2]. We were able to improve content, user interface and usability of the system based on the result of the sentiment analysis.

2 Related Work

The impulse for application sentiment analysis was the change of web standards—WEB 2.0. Since 2004, the world of the internet is not just about static websites, but users are already actively involved in the creation of content, and websites are full of interactive elements. The internet is full of subjective texts that can be further processed and analyzed for sentiment to gain valuable information [3].

Sentiment analysis gathers emotions of the author of the text. In its simplest form, the sentiment distinguishes positive and negative emotions, but there are also algorithms that are able to recognize fear, anger and other human emotions. Subjective texts, with identifiable emotions, are analyzed more frequently. There is also sentiment analysis of objective text based on the facts—for example monitoring of the financial market.

In general, sentiment analysis has been investigated mainly at three levels [4]:

Document Level: Result is the identification of positive, negative or neutral sentimnetu for the entire document. It is assumed, that each document contains text related to only one entity.

Sentence level: At this level, the task is to determine whether each sentence expresses positive, negative, or neutral opinion.

Entity and Aspect level: The document level and the sentence level analyzes do not exactly determine what people liked and disliked. Sentiment analysis at the aspect level is based on the rule that each opinion consists of sentiment (positive, negative) and objects (target of opinion).

One of the main conditions for determining the correct sentiment by the algorithm based on key words is quality of the lexicon. Lexicon contains sentiment words, also called opinion words, and multiplication words (almost, so, really etc.).

It is important to mention, that the sentiment analysis is a Natural Language Processing (NLP) problem. Success of the analysis also depends on quality of NLP for the chosen language, which is Czech in this case. There are many NLP methods and quality dictionaries for English. However, finding the right NLP techniques and vocabulary for other languages is much more complicated [5, 6].

Sentiment analysis in the Czech text is still a relatively unexplored area. When we were looking for the existing solutions, we found a work from 2011, where instead of dictionary, machine learning is used [7]. One of the most popular methods of machine learning is Support Vector Machine (SVM) [8]. The success of this method depends directly on the quality of the training set, which may be specific for different areas of entity [9].

There is also a hybrid method that combines the benefits of dictionary and machine learning approaches [10]. This article focuses on sentiment analysis of aspects based on the use of the new lexicon.

3 Sentiment Analysis

There are many emotions, but we only distinguish between positive, neutral and negative emotions for the resulting sentiment. We also divide the power of emotion into two groups—normal and strong. Special group are sentences which contain vulgar phrases and a negative sentiment is used. Table 1 shows all the defined classes of sentiment and also an example.

In the first step, the text is converted to the lower case format, and the system detects whether diacritics is used. If the diacritics is used, the system will use the original comparison in the next steps. Then the text is split to sentences using special characters and stopwords. In this context, sentence means a unit of sentiment analysis.

Each unit (sentence) is going through lemmatization - the words and expressions are converted to the basic word form by using the rules for inflection or the special dictionary. This problem is language specific, so the author's sense for the Czech language was used, and also the methods from [11]. Once we have a sentence in basic form, we can define n-grams (unigrams, bigrams, trigrams) [12]. With n-grams,

Table 1 Emotion classes

Emotion class	Description	Example
Vulgar	Vulgar and negative emot	"Buying this shit was a huge mistake!"
Negative2	Strong negative emotions	"I hate this camera, I am so angry about that"
Negative	Normal negative emotions	"This camera is not good"
Neutral	Neutral emotions	"I found some issues but it is not a bad camera"
Positive	Normal positive emotions	"I can recommend this camera"
Positive2	Strong positive emotions	"I love this camera I am so happy!"

Table 2 Formula transformation

Input	Output
"This camera has really amazing zoom"	{3} {M1.8} {POS2} {A123}
Explanation: {3}—three unrecognized words in row, "really"—{M1.8}(multiple word with cofficient), "amazing"—{POS2}(StrongPositive), "zoom"—{A123} aspect with ID 123	

accuracy of the algorithm is improved, because a comparison is done with words and their dependencies, not just each one separate word.

The main phase of the Sentiment Analysis is Formula Transformation. In this step, the application transforming text into stream of symbols. All n-grams are processed with the exception of dictionary, keywords dictionary, aspect list and emoticon list. The output is a formula, which is used for the final classification of the emotion class and relation with aspect. Table 2 describes the sentence-to-formula transformation.

The last part of the sentiment analysis is sentiment classification. In this step, evaluation and determination of the final sentiment of the text or aspect is done. For the example above is the result: Aspect—zoom, Sentiment—strong positive, Points +4.

3.1 Lexicon Based Acquisition

A good lexicon is the base of all methods of the lexicon sentiment analysis. The dictionary contains several types of phrases that have the main significance in the sentiment analysis [13], see Table 3.

For the creation of the dictionary, several automatic and manual methods can be used. One of the interesting automatic methods is based on the Wordnet lexicon [14], which contains the relations between words and is described in [15]. Since Czech is a very varied language, a semi-automatic method was used for generating the dictionary. This provides control over the phrases. At the beginning, we defined several

Table 3 Keywords types

Type	Description
Sentiment phrases	Phrases which identify positive or negative sentiment
Multiple phrases	Phrases which identify multiplication of the sentiment
Aspects	Attributes/properties of the subject
Exceptions	Phrases with special meaning

different phrases. The application automaticly gathers synonyms and semantically similar phrases in freely available dictionaries of synonyms of the Czech language. These dictionaries are based on Thesaurus. [1] Similar phrases were automatically proposed with the level of the sentiment related to the phrase or the same level [16].

Because lexicon is a very significant for our method , the new phrases were checked and others were added manually. With a sense for the Czech dictionary, it was possible to extend the exceptional phrases, which include the more complicated phrases. There are also special phrases which can significantly affect the final sentiment.

3.2 Aspect Based Analysis

Aspect-based sentiment analysis is a research problem that focuses on the recognition of all sentiment expressions within a given document and the aspects to which they refer [1]. Aspect-based sentiment analysis helps gather the right information from the text. It is good to know whether the customer feels positively or negatively about a product, but by using aspects, we are able to determine what exactly is good or bad about a particular product. This information is obviously valuable.

For a definition of aspects in the e-learning domain, we used two methods. One of them is that the aspect is a key word from the ontology of the learning content text. The second is that aspects are defined manually for the entire system, for example: font, code, navigation panel. There is also a possibility to use implicit phrases, which are stored as aspect synonyms. Example of implicit aspect phrase is “This camera is too heavy”, where we can recognize, that the author speak negatively about the camera weight. The word “heavy” is saved in the database of exception phrases and has a relation with aspect and also negative sentiment.

Aspect list can be extended also by using the automatic method. When the sentiment is recognized, the algorithm attempts to find a subject of the sentiment’s words. The subject is stored in a special database. If the subjects repeat in the text often, it becomes a candidate for a new aspect. For example, this way we extended aspects list of the e-learning system by adding the aspect “table” [17].

We can store a part of the analyzed text into the same database where the results of sentiment analysis are stored. When we also save a relation to the aspect and to

[1] http://thesaurus.com/

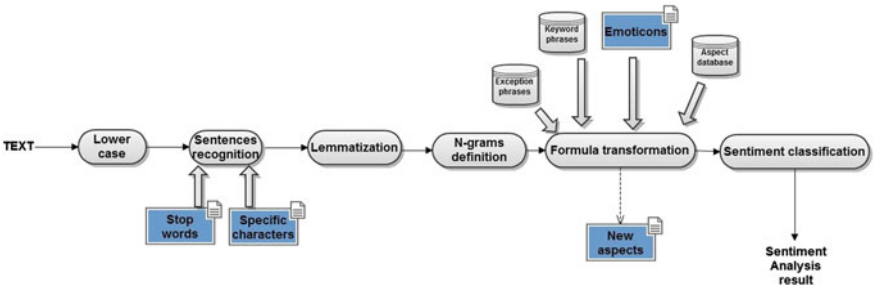


Fig. 1 The proposed system for sentiment analysis

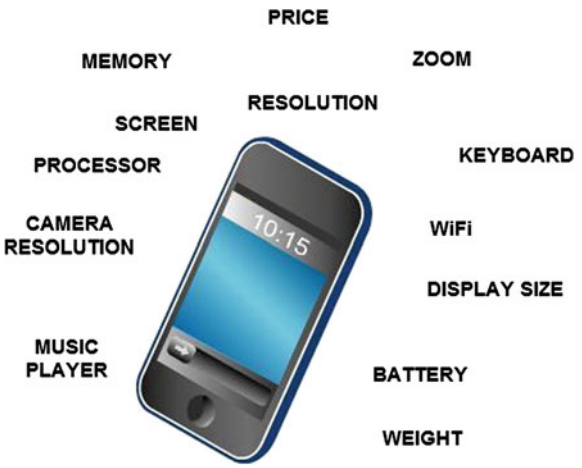


Fig. 2 Aspects of cell phone

the sentiment class we have a quality information. With this database is easy to find all comments or part of analyzed text for the specific sentiment and aspect.

For example we can get a list of sentences with strongly negative sentiment about the camera zoom. In these sentences can be found what exactly is wrong with the camera zoom. Result of the next analysis is what needs to be done for the customers satisfaction. But this analysis is based on small details in the text so manual human sentiment analysis is the best next step.

4 Experiment Results

As our experimenting environment we used elearning system XAPOS[2] [18]. Almost 200 of students add to the system 1473 comments. We defined more than 20 aspects and other was set by ontology. After the sentiment analysis algorithm was applied,

[2] http://arg.vsb.cz/XAPOS/, authors: Zdenek Velart, Petr Šaloun.

Table 4 Part of the output of XAPOS analysis

Aspect	Sentiment analysis
Example	4 negative2, 6 negative, 1 positive2
Table	11 negative, 4 neutral, 6 positive
Text	2 negative2, 4 negative, 5 neutral, 31 positive, 2 positive2

Table 5 Success of the algorithm—F Score

Type	Success(%)
Vulgar	100
Negative2	82
Negative	72
Neutral	82
Positive	74
Positive2	81

we got interesting picture about the system. In table below you can see a part of the output form the sentiment analysis of the student comments.

From the table above we know, that we should work on navigation and change the tables. On the other hand students like text of the learning objects in the system. We also successfully applied the algorithm in eshop domain.

We compared the results of the sentiment analysis of random comments from the users in e-learning and eshop domain with our manual sentiment analysis of these comments. The succcess of the algorithm was high. F score is a performance measure that combines precision and recall and ranges between 0 (worst performance) and 100 (best performance) [19]. It is mportant to say, that there is no absolute percentage of success, because sentiment analysis is subjective for each person opinion. See Table 5. for exact numbers.:

5 Conclusion and Future Work

Our contribution is sentiment analysis system for Czech language in real working system based on the lexicon method. We have demonstrated the usability of the algorithm in the domain of e-shops, and also in domain of e-learning. Data obtained from sentiment combined with the aspect relation were transferred to valuable information, which can help improve the product or system in the educational domain.

Value of the F-score 82 %, which explain the success of the algorithm, showed us, that the lexicon and whole method are not perfect and we need to work on the enhancement. For future work it is worth to consider using Wordnet or expand the lexicon by adding more n-grams and exceptions. Another option is the addition of AutoCorrection. There are also some challenges for the future work, such as irony and use of diacritics with only certain words. We believe that with these enhancements we can reach a big improvement of the algorithm for sentiment analysis of texts in the Czech language.

Acknowledgments The following grants are acknowledged for the financial support provided for this research: Grant Agency of the Czech Republic—GACR P103/13/08195S, by the Development of human resources in research and development of latest soft computing methods and their application in practice project, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational Programme Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic, and by Grant of SGS No. SP2013/114, VB—Technical University of Ostrava, Czech Republic.

References

1. Feldman, R.: Techniques and applications for sentiment analysis. Commun. ACM **56**(4), 82–89 (2013)
2. Krištofic, A., Bieliková, M.: Improving adaptation in web-based educational hypermedia by means of knowledge discovery. In: Proceedings of HT 2005 Sixteenth ACM Conference on Hypertext and Hypermedia, pp. 184–192. ACM Press. Sept. 2005
3. O'Reilly, T., Battelle, J.: Web 2.0 Five Years On, WEB2.0 SUMMIT—Special Report, pp. 15. O'Reilly Media Inc, (2009)
4. Liu, B.: Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, San Rafael (2012)
5. Neviarouskaya, A., Prendinger, H., Ishizuka, M.: Semantically distinct verb classes involved in sentiment analysis. IADIS Int. Conf. Appl. Comput. **2009**, 27–34 (2009)
6. Indurkhya, N., Damerau, F.J.: Handbook of Natural Language Processing, Second Edition, p. 704 (2010)
7. Cervenec, R., Burget, R.: Identifying expression of emotions in Czech text using semantic relations for dimension reduction. Elektrorevue **2**(3), 16–21 (2011)
8. Pang, B., Lee, L.: A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the Association for Computational Linguistics (2004).
9. Mullen, T., Collier, N.: Sentiment analysis using support vector machines with diverse information. National Institute of Informatics, Tokyo (2004)
10. Prabowo, R., Thelwall, M.: Sentiment Analysis—A Combined Approach. School of Computing and Information Technology, pp. 21. University of Wolverhampton (2009)
11. Sanda, P.: Determination of Basic form of Words, p. 68. Brno University of Technology (2011)
12. Hartmann, T., Klenk, S., Burkovski, A., Heidemann, G.: Sentiment Detection with Character n-Grams, p. 5. University of Stuttgart (2008)
13. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., Stede, M.: Lexicon-based methods for sentiment analysis. Assoc. Comput. Linguist. **37**(2), 267–308 (2011)
14. Fellbaum, C.: Wordnet: An Electronic Lexical Database, p. 423. MIT Press, Cambridge (1998)
15. Kamps, J., Marx, M., Mokken, R.J., de Rijke, M.: Using WordNet to measure semantic orientation of adjectives, Language and Inference Technology Group ILLC, p. 4. University of Amsterdam (2004)
16. Shanahan, J.G., Qu, Y., Wiebe, J.: Computing Attitude and Affect in Text: Theory and Applications (2006)
17. Wu, Y., Zhang, Q., Huang, X., WuPhrase, L.: Dependency parsing for opinion mining. Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pp. 1533–1541 (2009)
18. Šaloun, P., Velart, Z., Nekula, J.: Towards automated navigation over multilingual content. Studies in Computational Intelligence, vol. 418, pp. 203–229. Springer, Heidelberg (2013)
19. Van Rijsbergen, C.J.: Information Retrieval (2nd ed.), p. 147. Butterworth-Heinemann, Newton (1979)

How is the Process Network Organized and When Does it Show Emergent Properties in a Forest Ecosystem?

Juyeol Yun, Minseok Kang, Sehee Kim, Jung Hwa Chun, Chun-Ho Cho and Joon Kim

Abstract Ecosystems are open, self-organizing systems and energy of different quantity and quality provides the stimulus for organization, enabling different processes to progress at different rates. Here, information acts internally within the system to constrain its behavior, which can also flow into the system from outside, thereby prompting the self-organizing processes. The interplay of environmental conditions, energy, matter, and information defines the context and constraints for the set of processes and structures that may emerge during self-organization. Using the KoFlux tower-based measurements of energy, water and CO 2 flux time series in 2008 in a temperate forest in Korea, we have evaluated statistical measures of characterizing the organization of the information flow in ecohydrological process networks in a forest ecosystem. Here, process network is a network of feedback loops and direction of flow of matter, energy and information between the different variables. The goal of this study is to understand how ecosystem organization changes in time, and identify and characterize network-scale emergent properties by quantifying the varying ecosystem states. Ecosystem integrity is preserved when its self-organizing processes are preserved. The inherent challenges associated with the

J. Yun · J. Kim (✉)
Complex Systems Science Laboratory, Department of Landscape Architecture and Rural Systems Engineering, Seoul National University, Seoul 151-921, Korea
e-mail: joon@snu.ac.kr

M. Kang
National Center for Agro Meterology, Seoul National University, Seoul 151-744, Korea

S. Kim · J. Kim
Interdisciplinary Program in Agricultural and Forest Meteorology, Seoul National University, Seoul 151-744, Korea

J. H. Chun
Division of Forest Conservation, Korea Forest Research Institute, Seoul 130-712, Korea

C. -H. Cho
National Institute of Meteorological Research, Korea Meteorological Administration, Seoul 156-720, Korea

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_31,

time series data and the potential use of this conceptual approach and statistical tools are discussed for sustainable ecosystem management.

Keywords Self-organization · Flux monitoring · Process network · Information flow · Network statistics · Sustainable management · Forest ecosystem

1 Introduction

Complex systems are systems in which large networks of components with no central control and simple rules of operation give rise to complex collective behavior, sophisticated information processing, and adaptation via learning or evolution [3]. Thus, the science underlying complex systems should focus not only on the concepts of energy and matters, but also on those of feedbacks, information, communication, and purpose. Ecosystems are complex, open, self-organizing systems and energy of different quantity and quality provides the stimulus for organization, enabling different processes to progress at different rates. The interplay of environmental conditions, energy, matter, and information defines the context and constraints for the set of processes and structures that may emerge during self-organization.

Ecohydrological and biogeochemical processes associated with energy, water and carbon cycles in complex forest ecosystems can be viewed as a network of processes of a wide range of scales involving various feedback loops. Understanding such networks of feedback loops for key ecosystems in monsoon Asia is of great value and concern for resilient ecological-societal systems. Following Ruddell and Kumar [4, 5], we examined the dependence between a series of biophysical and meteorological variables measured at the flux tower in KoFlux by quantifying the information flow between the different variables along with the associated time lag. We tested the applicability of information theory to ecohydrological and biogeochemical systems with the datasets obtained at a temperate forest site in monsoon East Asia. The goal of this study is to understand how organization changes in time, and identify and characterize network-scale emergent properties by quantifying the varying system state.

2 Methods and Materials

2.1 Mutual Information and Transfer Entropy

We used Shannon's information entropy (H) as our methodology [7] and calculated the mutual information (I) and also the transfer entropy (T) to measure the reduction in the entropy of the current state of a measured variable Y_t due to the knowledge of prior state in another variable X_t, which is in addition to the information provided by the immediate prior history of X_t [1, 4]:

$$H(X_t) = -\sum p(x_t)\log p(x_t); \qquad 0 \le H(X_t) \le \log(m) \tag{1}$$

$$I(X_t, Y_t) = \sum_{x_t, y_t} p(x_t, y_t)\log\frac{p(x_t, y_t)}{p(x_t)p(y_t)} \tag{2}$$

$$T(X_t > Y_t, \tau) = \sum_{y_t, y_{t-\Delta t}, x_{t-\tau\Delta t}} p(y_t, y_{t-\Delta t}, x_{t-\tau\Delta t}) \cdot \log\frac{p(y_t\,|(y_{t-\Delta t}, x_{t-\tau\Delta t}))}{p(y_t\,|y_{t-\Delta t})} \tag{3}$$

This transfer entropy is computed from the component Shannon entropies following Ruddell and Kumar [4],

$$T(X_t > Y_t, \tau) = H(X_{t-\tau\Delta t}, Y_{t-\Delta t}) + H(Y_t, Y_{t-\Delta t}) - H(Y_{t-\Delta t}) - H(X_{t-\tau\Delta t}, Y_t, Y_{t-\Delta t}) \tag{4}$$

We normalized T using m (set at 11) discrete bins to estimate the probability distribution function. The information flow process network consists of the asymmetric pair wise T between the ith and jth variable from the set of n_v observed variables and is represented as an adjacency matrix.

The synchronization ratio, Tz, was calculated as the ratio of transfer entropy to the zero-lag mutual information, where

$$Tz(X_t > Y_t, \tau) = \frac{T(X_t > Y_t, \tau)}{I(X_t, Y_t)} \tag{5}$$

This ratio is used to characterize the nature of coupling between the dynamics identified through the time series. The term 'synchronization' is used as an intuitive approximation for the concept of mutual information, and the term 'forcing' to approximate that of transfer entropy. These couplings occur in pairs between each pair of variables, such that the coupling in one direction takes one type and time scale, and that in the other direction takes an independent type and time scale. Four types of couplings can be identified: (1) synchronization dominated (significant shared information but no significant information flow), (2) feedback dominated (significant information flow greater than significant shared information), (3) forcing dominated (significant shared information greater than significant information flow), and (4) uncoupled (no significant information flow or shared information) [4].

2.2 Study Site and Data Description

We used the time series data obtained in 2008 from the KoFlux forest tower site (i.e., GDK site) located in central Korea. The 30 min averages of eddy fluxes of energy, water and CO_2 were measured at 40 m above an oak-dominated old deciduous forest along with other micrometeorological variables. The description of the site and the data can be found in Kwon et al. [2] and AsiaFlux homepage (http://www.asiaflux.net). In this analysis, we selected 10 variables associated with

Table 1 Network matrix for mutual information for July (*top*) and August (*bottom*)

AI	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	63.1	41	7.3	5.5	9.4	1.2	7.1	3	3.3	3.7
GPP	41	60.9	8.1	8	12.6	1.2	4.5	2.4	2.8	4.5
H	7.3	8.1	61.3	8.3	17.6	x	5.7	5.6	4.1	5.7
LE	5.5	8	8.3	62	17.6	x	8.3	4.3	5.7	9.3
Rg	9.4	12.6	17.6	17.6	58.3	1.2	10.5	5.6	7	11.7
Precip	1.2	1.2	x	x	1.2	22.4	1.7	1.6	1.7	1.4
Re	7.1	4.5	5.7	8.3	10.5	1.7	92	11.4	29.2	22.7
Pa	3	2.4	5.6	4.3	5.6	1.6	11.4	93.8	10.5	7.5
T	3.3	2.8	4.1	5.7	7	1.7	29.2	10.5	90.9	24.2
VPD	3.7	4.5	5.7	9.3	11.7	1.4	22.7	7.5	24.2	80.9
AI	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	56.8	43.8	6.2	8.2	10.3	x	2.5	3.6	2.6	2.4
GPP	43.8	56	6.5	8.3	10.4	x	2.7	3.5	2.8	3
H	6.2	6.5	64	6.7	17.2	1.1	5.1	4.3	3.6	4.6
LE	8.2	8.3	6.7	61.5	13.7	x	4.2	4.6	2.9	5.9
Rg	10.3	10.4	17.2	13.7	56.2	0.9	5.9	6	4.7	8.1
Precip	x	x	1.1	x	0.9	19.8	2.7	2.1	3.1	2.1
Re	2.5	2.7	5.1	4.2	5.9	2.7	83.7	15	26.6	20.1
Pa	3.6	3.5	4.3	4.6	6	2.1	15	80.2	7.2	8.6
T	2.6	2.8	3.6	2.9	4.7	3.1	26.6	7.2	81.7	17.4
VPD	2.4	3	4.6	5.9	8.1	2.1	20.1	8.6	17.4	76.9

ecohydrological and biogeochemical processes in forests, which are atmospheric pressure (P_a), net ecosystem CO_2 exchange (*NEE*), gross primary productivity (*GPP*), ecosystem respiration (*RE*), latent heat flux (*LE*), precipitation (*Precip*), solar radiation (R_g), air temperature (T_a), vapor pressure deficit (*VPD*), and sensible heat flux (*H*). We computed process networks for each of thirty-six sub-daily time lags between 30 min and 18 h. Our spectral analysis shows that this subdaily time scale explained more than 30 % of the variances of the above variables associated with carbon and water cycles, reflecting that this range is an important scale of land-atmosphere interactions.

3 Results and Discussion

The monthly adjacency matrix for the 10 variables results in potential pairwise couplings, about 20–30 % out of which were found to be statistically significant at one or more time lags. The exemplary results on network matrix are presented for July and August of 2008 in Tables 1, 2, 3, 4.

Table 1 shows the matrix for the mutual information between pairs of variables at zero time lag. Source variable X index i is in rows; sink variable Y index j is

Table 2 Network matrix for uncertainty percentage for July (*top*) and August (*bottom*)

AIr	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	100	67.4	11.9	8.9	16.2	5.4	7.7	3.2	3.6	4.5
GPP	65	100	13.2	12.9	21.7	5.2	4.8	2.5	3.1	5.6
H	11.6	13.3	100	13.5	30.1	x	6.2	6	4.5	7
LE	8.7	13.2	13.6	100	30.2	x	9	4.6	6.3	11.5
Rg	14.9	20.8	28.6	28.4	100	5.5	11.4	6	7.8	14.5
Precip	1.9	1.9	x	x	2.1	100	1.9	1.7	1.9	1.7
Re	11.3	7.3	9.4	13.3	18.1	7.8	100	12.2	32.1	28.1
Pa	4.8	3.9	9.1	7	9.6	7	12.4	100	11.6	9.3
T	5.2	4.6	6.6	9.2	12.1	7.7	31.7	11.2	100	29.9
VPD	5.8	7.4	9.3	15.1	20.1	6.2	24.7	8	26.6	100
AIr	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	100	78.2	9.7	13.3	18.3	x	3	4.5	3.2	3.2
GPP	77.2	100	10.1	13.6	18.5	x	3.2	4.3	3.4	3.8
H	10.9	11.5	100	10.8	30.6	5.6	6.1	5.3	4.4	6
LE	14.4	14.9	10.4	100	24.3	x	5	5.7	3.5	7.7
Rg	18.1	18.5	26.9	22.2	100	4.7	7	7.5	5.7	10.6
Precip	x	x	1.7	x	1.7	100	3.2	2.7	3.8	2.7
Re	4.4	4.7	7.9	6.8	10.5	13.5	100	18.6	32.5	26.1
Pa	6.4	6.2	6.7	7.5	10.8	10.8	17.9	100	8.8	11.2
T	4.6	5	5.6	4.7	8.3	15.8	31.8	9	100	22.7
VPD	4.3	5.3	7.2	9.6	14.4	10.5	24	10.7	21.3	100

in columns. Matrix is symmetric. Italics indicate matrix diagonal. All values are in percent. Table 2 shows the matrix for the percentage of uncertainty of each Y explained by X.

Table 3 shows the matrix for the ratio of the maximum lag to mutual information for all significant couplings. Table 4 shows time lags of significant information flow on the interval, including the first significant lag, last significant lag, number of significant lags, and peak time lag. Significant lag times are [first-last (number), max].

The important goal of our study is to define the state of the forest ecosystem as a hierarchical pattern of coupling and feedback between subsystems. Subsystems are defined as a group of variables which are structurally equivalent such that they share a common role in the larger system structure. They are aggregations of individual nodes that share similar patterns of coupling type and time scale but hierarchy is allowed. Nodes in the same subsystem should share type (1) synchronization-dominated or type (2) feedback-dominated couplings. Type (3) forcing-dominated or type (4) uncoupled couplings mean that coupled nodes do not belong to the same subsystem [4].

The careful examination of Tables 1, 2, 3, 4 indicates that all the variables fall into three different subgroups. The first subgroup is the "synoptic" subgroup of T_a, VPD, P_a, and RE which are known to be strongly linked to the synoptic-scale

Table 3 Network matrix for the synchronization ratio of the maximum lag to mutual information for July (*top*) and August (*bottom*)

Tz(=T/I)	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	0.13	0.17	x	1.25	x	1.46	x	x	x	x
GPP	0.22	0.14	0.65	1.01	x	1.5	x	x	x	x
H	0.95	0.91	0.08	0.85	x	1.87	x	x	x	x
LE	1.47	1.04	0.66	0.11	x	2.07	x	x	x	x
Rg	1.05	0.74	0.34	0.51	x	1.64	x	x	x	x
Precip	x	x	x	x	x	0.08	x	x	x	x
Re	1.33	1.99	x	x	x	1.37	x	x	x	x
Pa	2.99	3.66	x	x	x	1.68	x	x	x	x
T	2.52	2.9	x	x	x	1.36	x	x	x	x
VPD	2.14	1.79	x	0.79	x	1.52	x	x	x	x
Tz(=T/I)	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	0.17	0.2	1.12	1.12	0.69	2.5	x	x	x	x
GPP	0.22	0.16	1.07	1.09	0.68	2.7	x	x	x	x
H	1.35	1.26	0.11	1.3	0.37	1.62	x	x	x	x
LE	1.1	1.07	0.99	0.14	0.52	3.06	x	x	x	x
Rg	1.09	1.06	0.4	0.69	0.12	2.02	x	x	x	x
Precip	5.46	5.15	2.35	4.33	2.68	0.07	x	x	x	x
Re	3.14	2.72	x	x	x	0.65	x	x	x	x
Pa	2.05	2.08	x	1.63	x	0.88	x	x	x	x
T	2.67	2.38	x	x	x	0.57	x	x	x	x
VPD	3.02	2.4	1.54	1.3	x	0.81	x	x	x	x

weather patterns. In this subgroup, ecosystem respiration, *RE*, is also included likely because it is a strong function of air and soil temperature. The second major subgroup is the "turbulent" subgroup associated with biophysical (e.g., energy balance) and biochemical (e.g., photosynthesis) processes on the land surface at plant canopy turbulent time scales. It includes *H*, *LE*, *NEE*, and *GPP*, which are variables whose information transport was found to peak typically at 0.5 to 1.5 h time scales. The third obvious subgroup is the "atmospheric boundary layer" (ABL) subgroup of *Rg*, and *Precip*, which are associated with ABL formation processes and convective activity.

By using the canonical couplings and the information in Tables 3 and 4, an arrangement of subsystems, information flow, feedback, and time scales can define the states of temperate forest ecosystem. Figure 1 shows the process network in two contrasting environmental conditions: (a) July (which is characterized by intensive, prolonged rainy season, i.e., *Jangma*) and (b) August (which is post-rainy season with high radiation and ample soil moisture). Typically, August and September are when the forest shows the second peak of net CO_2 uptake after *Jangma* associated with the Asian summer monsoon (e.g., Kwon et al. [2]).

In July, a synoptic subsystem was formed from T_a, *VPD*, P_a, and *RE* which shares synchronization-dominated coupling at very short time scales. However, ABL subsystem was less obvious and loosely organized with R_g and *Precip*, which shows

Table 4 Time lags of significant information flow on the interval for July (*top*) and August (*bottom*)

tau_timelag	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	1-17(10),1	1-6(11),1	x	0.5-3.5(7),0.5	x	18-18(1),18	x	x	x	x
GPP	1-17(13),1	1-11(13),1	1-3.5(2),1	0.5-5(8),0.5	x	13.5-17.5(2),13.5	x	x	x	x
H	0.5-18(13),2.5	0.5-18(14),2.5	1-1(1),1	0.5-3.5(7),0.5	x	6.5-14(4),13.5	x	x	x	x
LE	0.5-18(11),0.5	0.5-18(14),1.5	1.5-3(2),1.5	1-3(5),1.5	x	7-15.5(6),13	x	x	x	x
Rg	0.5-18(12),0.5	0.5-18(13),1.5	1.5-2(2),2	0.5-3.5(6),0.5	x	12.5-15.5(5),12.5	x	x	x	x
Precip	x	x	x	x	x	1-17.5(34),4	x	x	x	x
Re	0.5-18(15),1.5	0.5-18(10),1.5	x	x	x	11.5-11.5(1),11.5	x	x	x	x
Pa	0.5-12.5(24),8	0.5-9.5(4),1	x	x	x	1.5-17(7),12	x	x	x	x
T	0.5-18(16),2	1-17.5(7),13.5	x	x	x	10-11(2),10	x	x	x	x
VPD	0.5-18(9),18	0.5-18(8),18	x	17.5-18(2),18	x	7.5-7.5(1),7.5	x	x	x	x
tau_timelag	NEE	GPP	H	LE	Rg	Precip	Re	Pa	T	VPD
NEE	1-17.5(13),1	1-7(11),1	0.5-4.5(8),2	0.5-5.5(11),0.5	0.5-3.5(7),1.5	11-18(10),13	x	x	x	x
GPP	1-7(12),1	1-11(11),1	0.5-4(8),1.5	0.5-5.5(11),0.5	0.5-3.5(7),0.5	11-18(10),13	x	x	x	x
H	0.5-6.5(13),2	0.5-6(12),1	1.5-4(6),1.5	0.5-5(10),0.5	1.5-3(3),1.5	13-18(9),15.5	x	x	x	x
LE	0.5-18(14),1	0.5-18(13),0.5	1-2.5(4),1.5	1-4(7),2	0.5-2.5(5),1	10.5-18(13),13.5	x	x	x	x
Rg	0.5-18(14),0.5	0.5-5.5(11),0.5	1-3(4),2.5	0.5-4.5(9),0.5	1-3(5),1.5	10-18(17),18	x	x	x	x
Precip	3-12.5(20),5	3-13.5(19),5	6-7(3),6.5	4.5-11(13),8	6-10(8),6.5	1-4.5(8),2.5	x	x	x	x
Re	0.5-18(22),1.5	0.5-18(22),1.5	x	x	x	1.5-12.5(5),5.5	x	x	x	x
Pa	0.5-18(29),7.5	0.5-16.5(17),1.5	x	2-2(1),2	x	0.5-4.5(7),0.5	x	x	x	x
T	0.5-18(27),0.5	0.5-18(27),10	x	x	x	3.5-12(2),3.5	x	x	x	x
VPD	0.5-18(28),18	0.5-18(24),17.5	18-18(1),18	16-18(5),17	x	10.5-14(8),11.5	x	x	x	x

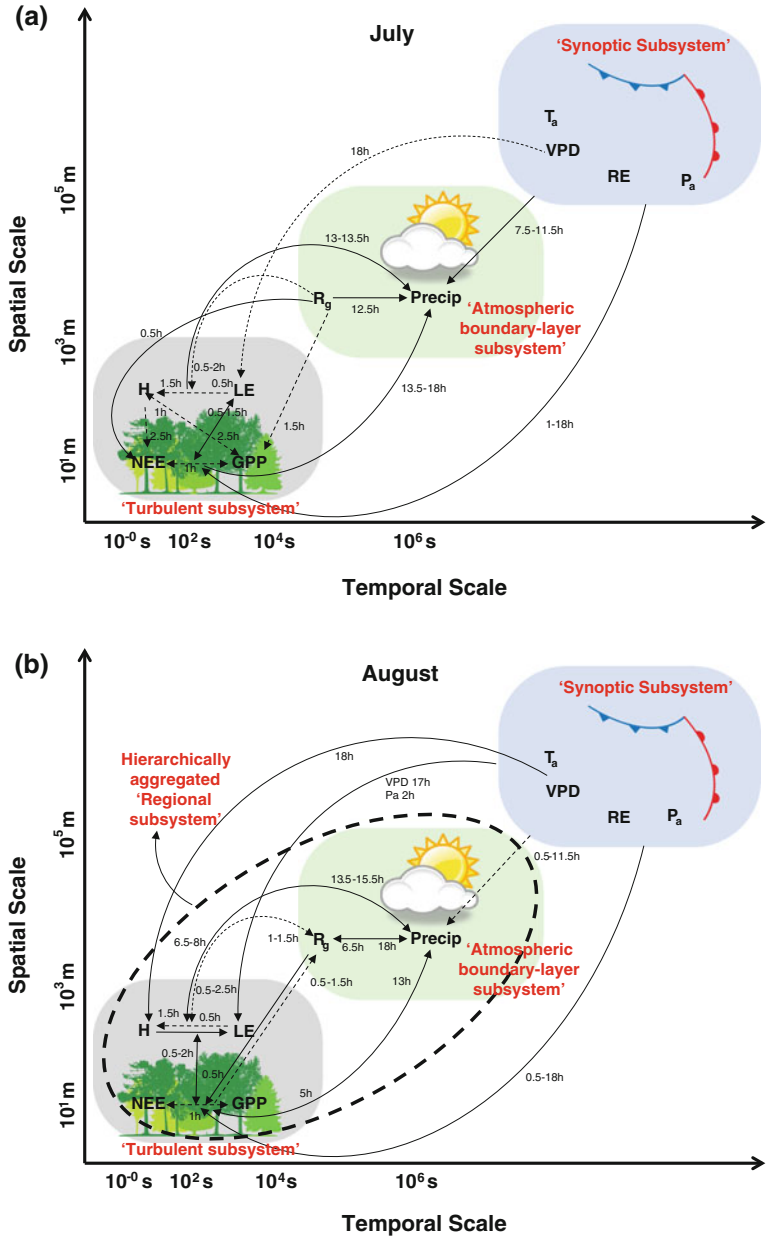


Fig. 1 The process networks in (**a**) July ('Jangma' - intensive rainy season) and (**b**) August (post-rainy season) in 2008 in Gwangneung temperate deciduous forest ecosystem in central Korea. Type (1), (2), (3), and (4) couplings result in the interpretation of this forest ecosystem as three subsystems linked at time scales ranging from 0.5 to 18 h. *Solid (dotted) arrow* represents forcing-dominated (feedback- dominated) couplings. *Solid (dotted) circle* denotes subsystem (aggregated subsystem). (The figure frame was adapted from Ruddell et al. [6])

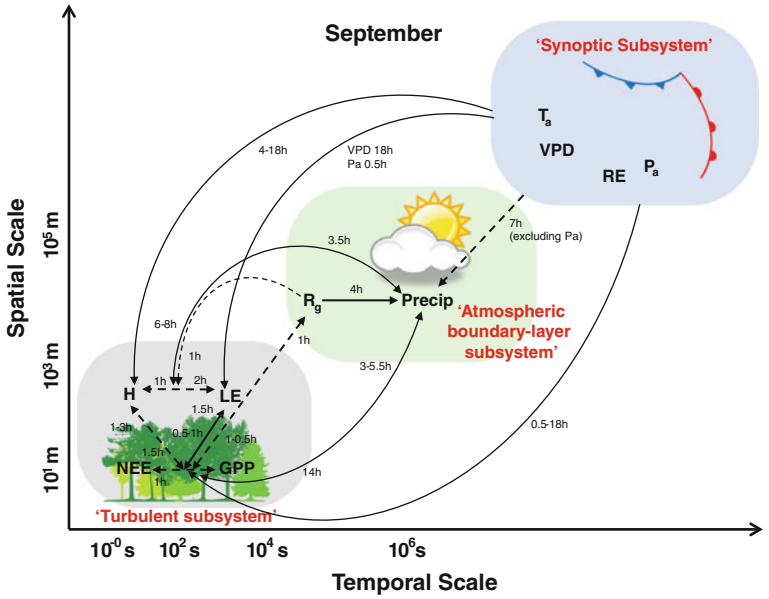


Fig. 2 The process network in September 2008 in Gwangneung temperate deciduous forest ecosystem in central Korea. (The data retrieval rate in September was very low due to power failure and instrument malfunction, resulting in 74 % gap-filling of the missing data.)

forcing-dominated connection at longer sub-daily time scale of 13 h. Turbulent subsystem was partially formed among the biochemical feedback loop (between *NEE* and *GPP* which share feedback-dominated couplings at the 1 h time scale), H and *LE* (which share both feedback- and forcing-dominated couplings at 1–2 h time scale). There was no indication of hierarchical aggregation for self-organization. The synoptic subsystem exported information to the other two subsystems whereas information also flew bottom up from the forest photosynthesis and energy fluxes. In July, precipitation was information sink.

In August, a fully self-organizing turbulent subsystem was formed between the biophysical feedback loop (between *H* and *LE*) and biochemical feedback loop (between *NEE* and *GPP*), which share forcing-dominated couplings at the 1–2 h time scale. The synoptic subsystem was information source for the other two, and served as a large-scale forcing on the turbulent subsystem and as a feedback to ABL subsystem. The turbulent and ABL subsystems were coupled via feedback loops, and form a regional self-organizing subsystem that is a hierarchical aggregate of those two subsystems. This regional-scale feedback subsystem forms the basis for forest ecosystem to self-organize at the regional scale through collective dynamics during the second peak of healthy growing season. In this state, information flew both top-down from synoptic weather systems and bottom-up from the forest photosynthetic activity and energy partitioning (e.g., Ruddell and Kumar [4, 5]).

Such an aggregate of regional subsystem was also observed during the peak growth stage in May when the forest's net carbon uptake showed the first peak. The aggregate of regional subsystem broke apart in June with the onset of summer monsoon (not shown). We expected to observe the process network in September to be similar to that observed in August. However, not only the aggregate of regional subsystem was absent from the process network (Fig. 2). Different characteristics of time lags, coupling types, and the absence of the regional subsystem are attributed to an artificial gap-filling of the missing field data in September due to power failure and instrument malfunction. These results demonstrate a potential risk of the use of gap-filled data which may inadvertently modify the system's inherent process network, resulting in inappropriate understanding of the processes.

4 Concluding Remarks

The central role of information in directing self-organizing processes and structures has only recently been put forward. One of the recent progresses in developing methodologies to describe such role is the information flow process network approach [4, 5]. A process network is defined as a network of feedback loops and the associated time series that depicts the magnitude and direction of flow of energy, matter and information between different variables. Our results support that a process network approach can be used to formally resolve feedback, time scales, and subsystems that define the complex ecosystem's organization by considering mutual information and transfer entropy simultaneously. We have shown that the turbulent and ABL subsystems are coupled through feedback loops, and form a regional self-organizing subsystem that is a hierarchical aggregate of those two subsystems in August when the forest is in healthy environment. Also noted is the disappearance of the self-organizing subsystem in the process network when the time series data were artificially gap-filled for missing data, which is a common practice in post-data processing of field observation data. Sustainable management is about maintaining the integrity of the combined ecological-societal systems. Integrity is preserved when the system's self-organizing processes are preserved, something that happens naturally if we maintain the context for self-organization in ecological systems, which, in turn, will maintain the context for the continued well-being of the societal systems. Further studies are in progress on the application of network statistics to measure the statistical feedback, entropy, and net and gross information production of subsystems on the ecosystem network to characterize network-scale emergent properties.

Acknowledgments This study was funded by the Korea Meteorological Administration Research and Development Program under Grant CATER 2012-3030. The dataset used for this study was processed through the support of Research Settlement Fund for the new faculty of Seoul National University. We thank Bindu Malla Thakuri and Jaeil Yoo for their field support and data analysis and Boeun Choi for the preparation of Figs. 1 and 2.

References

1. Kumar, P., Ruddell, B.L.: Information driven ecohydrologic self-organization. Entropy **12**, 2085–2096 (2010)
2. Kwon, H., Kim, J., Hong, J., Lim, J.-H.: Influence of the Asian monsoon on net ecosystem carbon exchange in two major ecosystems in Korea. Biogeosciences **7**, 1493–1504 (2010)
3. Mitchell, M.: Complexity: A Guided Tour. Oxford University Press, New York (2009)
4. Ruddell, B.L., Kumar, P.: Ecohydrologic process networks: 1. Identification. Water Resour. Res. **45**(W03419), 2008W (2009). doi:10.1029/R007279,2009
5. Ruddell, B.L., Kumar, P.: Ecohydrologic process networks: 2. Analysis and characterization. Water Resour. Res. **45**(W03420), 2008W (2009). doi:10.1029/R007280,2009
6. Ruddell, B.L., Brunsell, N.A., Stoy, P.: Applying information theory in the geosciences to quantify process uncertainty, feedback, scale. Eos **94**(5), 29 (2013)
7. Shannon, C.E.: A mathematical theory of communication, Bell Syst. Tech. J. **27**, 379–423 (1948)

Part IV
Complex Systems Science Applications

Active Control Metrology for Preventing Induced Thermal Damage During Atmospheric Pressure Plasma Processing of Thermal Sensitive Materials

Victor J. Law and Denis P. Dowling

Abstract The successful surface activation and joint promotion of lightweight aircraft grade carbon composites using atmospheric pressure plasma jets has the potential to improve the economics of the aircraft industry. To achieve these economic savings the technological challenge of plasma processing of thermally sensitive composites must be met in an industrial manufacturing environment. This chapter describes an acoustic plasma control strategy that is based upon the complex acoustic dynamics of the plasma-composite interaction.

Keywords Plasma · Thermal damage · Composite · Acoustic · Process control

1 Introduction

Atmospheric pressure plasma jets (APPJs) in the form of needles, pens and pencils have the unique ability to produce cold plasma that contain a range of ions and activated molecular species that can be used in the treatment of skin with applications ranging from wound healing to anti-wrinkle treatments. Their use is also recognised in the repair of carbon composites and electrical wiring loops. The fact that these plasma devices can be hand-held and manually directed has led to safe and easy to handle designs that have fixed, or limited variability, power units and gas supply systems. It has been shown for static dwell times of a second and gap distances of a few 10 s of mm, composite surfaces undergoes varying degrees of modification ranging from surface activation to plasma induced damage. The latter may be categorised into three broad and overlaying morphological types: (a) surface charring, (b) resin reflow and, (c) resin ablation leaving the carbon fibres mechanically

V. J. Law (✉) · D. P. Dowling
School Mechanical and Materials Engineering, University College Dublin, Belfield, Dublin 4, Ireland
e-mail: viclaw66@gmail.com

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_32,

intact [1]. Engineering trials are on-going with the knowledge that there is a need to fully understand the plasma device and how plasma alters both thermally sensitive living tissue and composites material.

To prevent the 'hot' visible plasma region coming in to direct contact with thermally sensitive materials hollow spacers (cages) are supplied to keep the 'hot' region at a safe working distance thus tacitly recognising the possibility of induced plasma thermal damage. Passive safety control features that are built in to the plasma device, such as nozzle cages, are but one solution, but as in many complex systems an active monitoring system that provides feedback control and early warning is desirable and necessary when the plasma process moves outside normal operational conditions.

This chapter presents an active control strategy that can be independently deployed to monitor the heterogeneous plasma-surface reaction zone without the need to access the plasma device electrical power source, or access the chemical information within the plasma-surface process. The measurement metrology comprises a single omnidirectional condenser microphone that is positioned perpendicular (90°) to the plasma for sensing the plasma nozzle-to-surface distance and hence induced surface temperature.

Including this introductory section, the chapter is dived into five sections. Section 2 details Tepla Plasma-Pen^{TM} APPJ and Sect. 3 the deployment of the single microphone to measure its compound nozzle far field acoustic emission pattern, and the analysis of the nozzles mode of operation. Section 4 describes real-time deconvolution of the plasma sound radiation in the frequency-domain as the plasma engages with a composite surface: the deconvolution components (A1 = selected peak amplitude, A2 = selected span amplitude and A3 = total averaged amplitude) [2]. These three components are then reconstruction in 3-dimensional state-space to represent the dynamics of the time evolving plasma-surface process. Similar plasma state-space reconstruction processes based on electrical data have been reported and can be found in references [3–6]. This state-space representation is then compared to the composite surface temperature and induced water contact angle. Section 5 provides the conclusion to this work.

2 Plasma-Pen

The PVA Tepla Plasma-Pen^{TM} is a miniature blown arc plasma source which uses dry compressed air as the working gas and to some degree can be viewed as miniature version of the PlasmaTreat Openair^{TM} system [2, 7]. In terms of polymer activation, the Pen produces a thermal power transfer of 100 s mWatts to the surface as compared to 1–5 W for the PlasmaTreat system. The size and weight of the Pen allows it to be used as a hand-held device or mounted on a CNC XYZ gantry for precise plasma treatment of materials. The Pen is driven by a 100 Hz, 1.5 kV Peak positive dc pulse repetitive chain sequence derived from a mains frequency of 50 Hz with full rectification without capacitance filtering or regulation. The nozzle head is located at the end of 3.0 m screened cable and is built from nickel plated copper with an

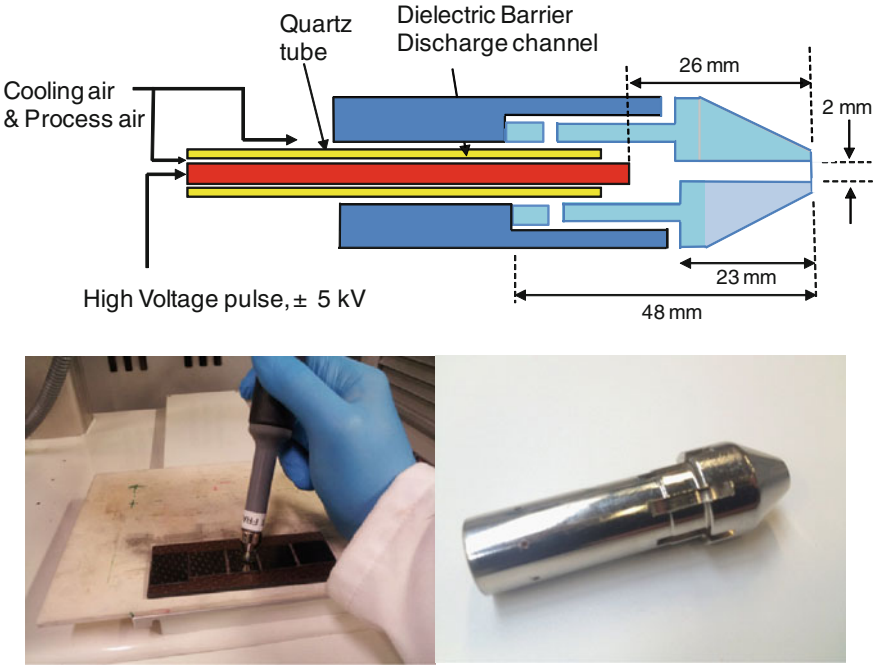


Fig. 1 Schematic and Photograph of Plasma-Pen and its Nozzle Head

on-axis 2 mm internal diameter nozzle with internal length of 23 mm from where the orifice diameter expands in to the cathode chamber. Typically 4.7 l/m compressed air is forced through the nozzle at a Marc number of 0.072 M. For cooling purposes the head is a compound design where a proportion (15 l/m) of air is extracted from the cathode region by six radially spaced 1 mm diameter ports which redirects the sampled air along the outer surface of the head to six radial equally spaced slots. The function of these slots is to provide a counter surface to increase conductive heat lost and redirect the air flow away from the nozzle exit port: similar to the action of a muzzle brake [8]. It is these design features which gives the Pen its perceived modulated audible emission. Visually the expelled plasma gas appears as a decaying luminous needle with a distal length of 1.5 cm. Infrared thermography temperature measurements of the plasma-composite reaction, indicates a composite surface temperature of approximately 140 °C at a nozzle-to-surface distance of 2 mm and which exponentially falls to 35 °C at a distance of 30 mm. The standard Plasma-Pen design does not provide accessible voltage and current probes for monitoring the electrical power supply, nor can the dried compressed air flow be varied. Given these limitations and the loudness of acoustic noise [SPL ∼60 dB], this APPJ is a suitable candidate for acoustic surface metrology. A schematic and two photographs detailing the APPJ head and manual application to a composite surface are shown in Fig. 1.

3 Plasma-Pen Far Field Acoustic Measurements

This section describes the use of a omnidirectional microphone to capture the APPJ far field acoustic emission pattern. This method of acoustic measurement is in contrast to normal practice within the aircraft industry where jet engine noise patterns are measured both in near field using microphones and in far field using a linear array of microphones and where the measurements are performed within an anechoic chamber [9, 10]. In ordered to capture the far field acoustic pattern, the microphone position is scaled to a distance of 40x the jet diameter (80 mm) and the directivity of the emission measured between 60 to 300°, where 90 and 270° are perpendicular to the nozzle and a 180° position that faces the gas flow emanating from the on-axis nozzle aperture.

The microphone signal was passed to a laptop computer soundcard and sampled at a rate of 44.1 kHz. The sampled signal was then processed with National Instrument LabVIEW using a Fast Fourier Transformation to display the sampled data in the frequency-domain from 0 to 22 kHz. Figure 2 shows the emission measurements angles: 90, 130 and 180°. The 90° measurement which is perpendicular to the jet nozzle reveals three discrete acoustic peaks in the range 3–8 kHz: with the asymmetric peak at 3.3 kHz having the greater amplitude and width as compared to the neighbouring peaks at 5.2 and 7.4 kHz. As the microphone is repositioned further along the jet and angled at 130° to the nozzle exit the acoustic peak emissions changes in amplitude and morphology with the peak at 3.3 kHz exhibit a falling in intensity (from −33 to −35 dB). At the 180° position the acoustic emission noise floor increases and approaches an inversely proportional to frequency dependency with a rate of $-1.7\,\mathrm{dB.kHz^{-1}}$. Under these conditions the fall in amplitude of the three discrete peaks allows them to be partly obscured by the noise floor.

For a constant air flow through the nozzle, the acoustic sound emission under plasma and non-plasma conditions has been made under free space conditions (nozzle 10 cm away from a surface) and the 0 and 12 kHz results are shown in Fig. 3. The comparative measurements reveal that air flow only generates a lower acoustic intensity, as compared to plasma, across the measurement frequency spectrum. In addition the discrete peaks have shifted up in frequency by approximately 0.5 to 1 kHz. The amplitude of the three peaks is also some 15 dB lower than the in plasma condition. A further feature of note is above 7 kHz the air acoustic intensity falls more rapidly with frequency.

Putting to one side the multiple peak signature for a moment, the enhanced amplitude in the plasma acoustic signal may be understood by considering the jet's Strouhal number (*St*) which is a dimension-less measure of how the drive frequency ($f_d = 100\,\mathrm{Hz}$) is synchronised to the air velocity ($v = 25\,\mathrm{m/s}$) through the nozzle orifice. The scale length of the nozzle is $D = 0.002\,\mathrm{m}$. These quantities are defined in Eq. (1). In this equation when $St \sim 1$, the drive frequency is synchronized with the air velocity with in the nozzle. For low *St* value, the quasi steady state of the gas dominates the oscillation. And at high values of *St* the viscosity of the gas dominates fluid flow ("fluid plug"). In our case the non-plasma air flow produces a low *St* value (0.008).

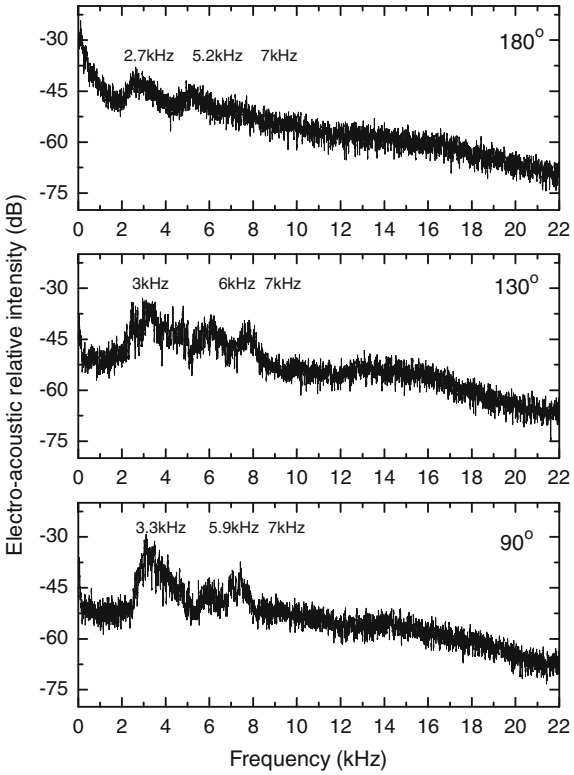


Fig. 2 Far field frequency dependant acoustic emission for three measurement angles (90, 130 and 180°)

$$St = \frac{f_d D}{v} = 0.008 \tag{1}$$

It is well established that the use of dielectric barrier discharge (DBD) actuators in aircraft jet engines [9–12] can alter the engine acoustic noise spectrum by varying DBD drive frequency. For example a low DBD drive frequency excitation that producing a $St \sim 0.5$ leads to a broadband increase in jet sound emission, while the high frequency excitation of a jet ($St > 2$) leads to broadband reduction of noise provided that the action is sufficiently intense above the air noise level. This frequency transformation in the acoustic energy spectrum is put to good use in engine noise mitigation [9–12] and the altering the air flow dynamic on aerodynamic surfaces [13] plasma fairing of undercarriage [14] and weapons bay on high speed platforms [15]. This mechanism has also been proposed as the source of acoustic amplification within plasma speakers [16] and the kINPen MED APPJ [17].

The production of electronic winds by the DBD may be put forward as one possible mechanism for this acoustic amplification, where the unsteady longitudinal electronic winds are composed of charged ions and move under the influence of

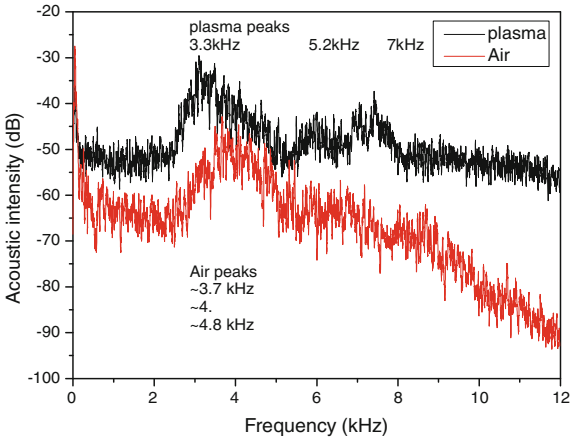


Fig. 3 Far field 90° to nozzle acoustic measurement of Plasma-Pen

the positive and negative going edges of the voltage pulse as applied by the power supply. Due to the high collisionality of charged particles and neutrals at atmospheric pressure, the charge particles quickly transfer all momentum energy to the neutrals so enabling a phased locked or quasi-synchronised conditions to be generated when $St < 1$. Electrons are not considered here due to their low mass. Under these condition the electronic winds drive/enhance the velocity of the neutral gas molecules traveling through the nozzle, to the vibrational antinode (nozzle exit) where air is free to undergo increasing alternating compression and rarefication which is perceived as an increased loudness in the radiated sound energy without affecting the position of the nozzle's fundamental resonant frequency.

From an aerodynamic point-of-view a jet sound source is fairly easy to model as sound energy being directly radiated into the environment from some opening. Using the measured sonic values and the known nozzle geometry a phenomenological mathematically model of the longitudinal acoustic emission mode is considered in this section. The first model to be considered is the closed air-column model, or clarinet model, in Eq. (2) [17, 18].

$$f_n \approx \frac{nc}{4\,(L + 0.6r)} \tag{2}$$

Where the modulo character (n) enables the overtone frequencies to be identified, otherwise L is the physical length of the nozzle (L = 23 mm), 0.6r is the end correction [19] where r is the internal radius of the nozzle (0.001 m), and c is the sound velocity in air (346 m.s^{-1} at 25 °C). The denominator value 4 describes the quarter-wavelength longitudinal mode. For the quarter-wave model the nozzle exit aperture defines the maximum pressure vibration (or acoustic antinode), whereas and the internal nozzle aperture is the acoustic node due to flow gas is being compressed with respect the nozzle exit. To confirm the nozzle temperature, infrared thermography of the head

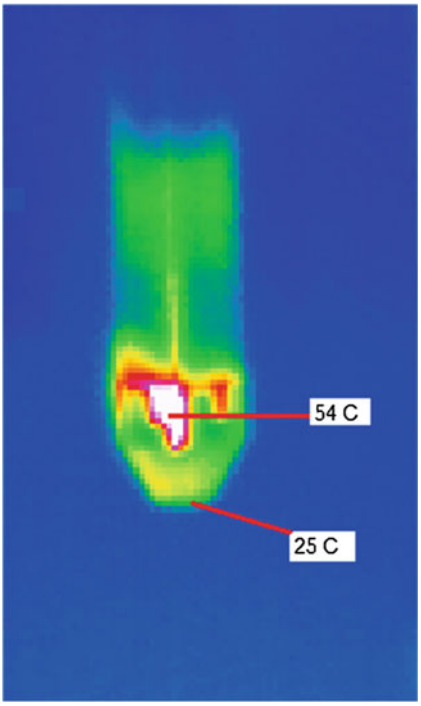


Fig. 4 Thermographic image of Plasma-Pen nozzle after 20 seconds of operation

after 20 s of operation has been performed, see Fig. 4. The measurement reveals that the nozzle temperature is generally 25 °C and the slots have a measured temperature of 54 °C. Now consider the plasma and air acoustic peaks in Fig. 3 with regard to the nozzle geometry and Eq. (2). Consider first the three discrete plasma acoustic peaks at $f_\circ \sim 3.3$, 5.2 and 7 kHz, and the known physical geometry of the nozzle ($L = 23$ mm), when evaluating Eq. (2) for $L + 0.6r$ with a sound velocity of 346 m.s^{-1}. The evaluations yields a value of $L + 0.6r = 26$ mm for 33 kHz and 16.6 and 12.3 mm for the remaining two higher frequencies. This simple computation suggests that the nozzle is performing as described by the quarter-wave equation with the lower discrete frequency being the nozzle natural resonance. Infrared thermography of the slots (not shown in abstract) reveal that the nozzle temperature is generally 10 °C hotter than the slots and both are considerably hotter than the surrounding metal.

For the same air flow rate, but with the plasma turned-off condition and a laboratory air temperature measured to be of the order 20 °C the nozzle resonance frequency would be expected to be slightly higher. Indeed, for $c = 343$ m.s^{-1}, the first discrete frequency is 3.7 kHz.

The far field radial acoustic component aligned to the cooling slots and mid points between the cooling slots reveal two contrasting acoustic signatures. The cooling slots produce a strong peak at 3 kHz, as in Fig. 2, and broadband acoustic radiation between 3 and 6 kHz. Infrared imaging also reveals that there is a local variation in

the heat pattern, with heat maxima at the cooling slots, being some 10 °C hotter than the mid points.

4 Plasma-Surface Process 3-Dimensional Mapping

With the fundamental plasma acoustic peak identified (∼3 kHz) principle component analysis (PCA) can be performed on the acoustic peak to locate and register the treatment surface [2]. The successful outcome of which, in this work, is coupled with the knowledge of the plasma induce temperature of the composite, water contact angle and type of induced damage. The PCA is performed using three acoustic performance parameters or descriptors from the frequency-domain signature to build a 3-dimensional state-space representation of the plasma process as a function plasma engagement with the composite surface. The performance parameters of the acoustic signal are as:

1. A1: The specific acoustic level at the fundamental acoustic frequency (3 kHz)
2. A2: Averaged specific acoustic emission over a 500 Hz span at a 3 kHz
3. A3: The mean electro-acoustic emission measured over the total 0–12 kHz span

These three parameters are mapped separately to one the three axis (A1 to X, A2 to Y, and A3 to Z) of a 3-dimensional graph were the vectors are updated continuously to produce a time evolving image of the dynamic heterogeneous plasma-composite reaction. The result of a typical mapping process is shown in Fig. 5. Here it can be seen that three dataset clusters are generated. The first cluster maps a 33–22 mm gap distance where composite surface temperature ranges between 30–40 °C. As the plasma engages closer to the composite surface (11 mm), the three parameters change in amplitude and form a new and separate cluster. At this point the composite surface temperature has increased to 55–65 °C. Finally with the gap distance set to 1 mm, an new cluster of greater average acoustic amplitude (A3) which is accompanied by an even higher composite surface temperature of 140 °C. Note the geometric centre of each cluster are shown in the insert to Fig. 5.

Water contact angle measurements reveal a strong correlation with the composite surface temperature within the dynamic state-space representation. Table 1 tabulates this correlation. Here it can be seen that at extended engagement distances 25–30 mm (composite temperatures of 30–50 °C) the water contact angle is little altered by the plasma process. At 10–11 mm distance (composite temperatures of 55–65 °C) the composite surface becomes activated with a water contact angle of ∼10°. Plasma engagement with the surface at 2 mm increases the surface temperature further to 140 °C with a water contact angle of 10–15°. However between 4 and 8 mm distance the water contact angle increases to 40–50° with a surface temperature of 80–90 °C. These last two regions are associated with plasma induced thermal damage of composite: charring, resin reflow, and resin ablation.

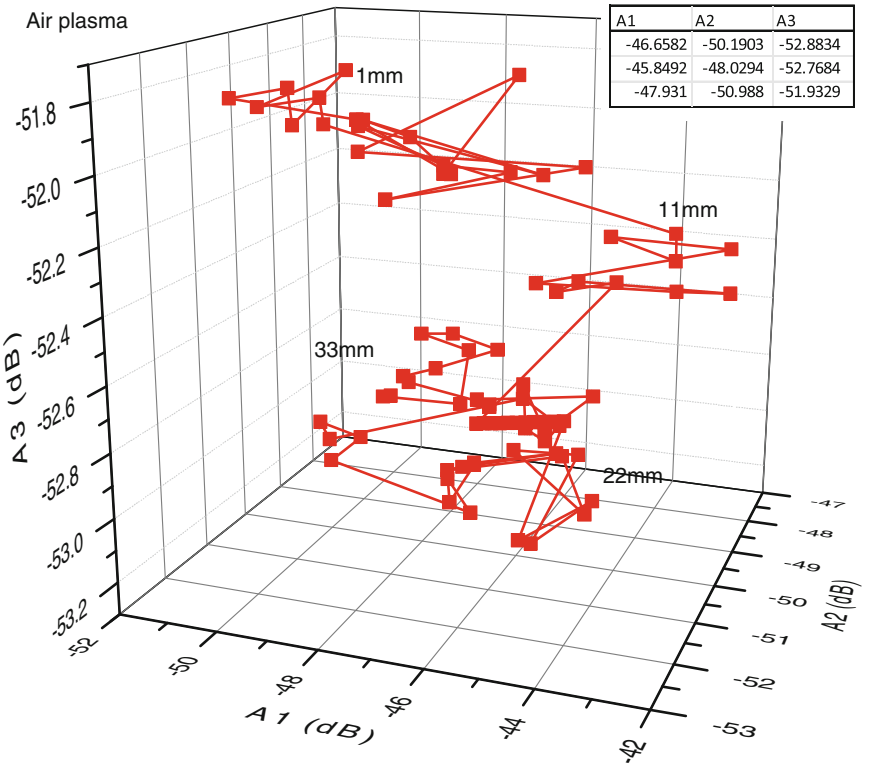


Fig. 5 Plasma-Pen acoustic state-space representation of engaging with a composite surface. The insert table depicts the geometric centre of each cluster

Table 1 Composite plasma induced thermal damage parameters

Distance	33–22 mm	11 mm	1 mm
Temperature (°C)	30–40	55–65	140
WCA (degrees)	70–75	10	40–10
Damage	No apparent reaction	Activation	Carring / reflow/ ablation

5 Conclusion

This work has presented an active control strategy for the prevention of plasma induced damage when aircraft grade composite materials are treated with an APPJ. The APPJ used for this study is the air operated PVA Tepla Plasma-Pen, where the control strategy is based upon far field acoustic response of the APPJ and its interaction with a composite surface. The thermodynamic data is obtained using infrared thermography imaging of the plasma-composite process. The APPJ acoustic

response is modelled using a quarter-wavelength closed end air column model (Eq. 2) and the outcome of which is used to identify, and separate from the APPJ, the plasma-composite interaction.

Using this spatial-temporal thermodynamic knowledge, the Plasma-Pen acoustic signature, as it engages with the composite can be mapped in real-time using 3-dimensional state-space to representation of the plasma-surface interaction. The resultant spatial-temporal 3D state-space can be used to map the composite surface temperature and the type of plasma induced damage. The mapping thus enables preventative measures to be taken before the composite undergoes irreversible plasma damage. Although this type of state-space representation has been used for plasma-electrical datasets it is thought that acoustic information coupled with surface temperature and water contact angle is novel and not been reported before, in particular for thermal sensitive composite materials.

Acknowledgments This work is partially funded by the ICOMP Composites Technology Centre programme.

References

1. Mohan, J., Ramamoorthy, A., Ivankovic, A., Dowling, D.P., Murphy. J.: Effect of an atmospheric pressure plasma treatment on the mode I fracture toughness of a co-cured composite joint. J Adhesion (Accepted 2013)
2. Law, V.J., O'Neill, F.T., Dowling, D.P.: Evaluation of the sensitivity of electro-acoustic measurements for process monitoring and control of an atmospheric pressure plasma jet system. PSST 20(3), 035024 (2011)
3. Law, V.J., Ramamoorthy, R., Dowling, D.P.: Real-time process monitoring during the plasma treatment of carbon weave composite materials. JMSE **1**(2B), 164–169 (2011)
4. Walsh, J.L., Iza, F., Janson, N.B., Law, V.J., Kong, M.G.: Three distinct modes in a cold atmospheric pressure plasma jet. J. Phys. D: Appl. Phys **43**(7), 075201 (14pp) (2010)
5. Law, V.J., Dowling, D.P., Walsh, J.L., Iza, F., Janson, N.B., Kong, M.G.: Decoding of atmospheric pressure plasma emission signals for process control. CMSIM J. **1**, 69–76 (2011)
6. Topala, I., Dumitrascu, N., Dimitru, D.-G.: Experimental and theoretical investigation of dielectric-barrier plasma jet in helium. IEEE PS **40**(11), 281–2816 (2012)
7. Dowling, D.P., O'Neill, F.T., Langlais, S.J., Law, V.J.: Influence of dc pulsed atmospheric pressure plasma jet processing conditions on polymer activation. PPP **8**(8), 718–727 (2011)
8. Hammer, E.W.: Muzzle Brake Theory. vol. 2, pp. 320. The Franklin Institute Research Laboratories, Philadelphia (1949)
9. Kopiev, V.F., Bityurin, V.A., Belyaev, I.V., Godin, S.M., Zaitsev M.Y., Klimov, A.I., Kopiev, V.A., Moralev, I.A., Ostrikov, N.N.: Jet noise control using the dielectric barrier discharge plasma actuators. Acoust. Phys. **58**(4), 434–441 (2012)
10. Samimy, M., Kim, J.-H., Kastner, J., Adamovich, I., Utkin, Y.: Active control of a Mach 0.9 Jet for noise mitigation using plasma actuators. AIAA J. **45**(4), 890–901 (2007)
11. Ginevsky, A.S., Vlasov, Y.V., Karavosov, R.K.: Acoustic control of jets. Babitsky V.I., Witterbug J. (eds.) Springer, Heidelberg (2004)
12. Moreau, E.: Airflow control by non-thermal plasma actuators. J. Phys. D: Appl. Phys. **40**, 605–636 (2007)
13. Gulec, A., Oksuz, L., Hershkowitz, N.: Optical studies of dielectric barrier plasma aerodynamic actuators. Plasma Sources Sci. Technol. **20**, 045019 (7pp) (2011)

14. Thomas, F.O., Kozlov, A., Corke, T.C.: Plasma actuators for bluff body flow control. AIAA 2006–2845, 1–16 (2006)
15. Schwimley, S.L., Foristell, M.O., Drouin-Jr, D.V., O'Fallon, I.L.: Plasma actuator method for use with a weapons bay on a high speed mobile platform. Eur. Pat. EP 1 995 172 B1 (Date of filling 16.05.2008)
16. Sutton, Y., Moore, J., Sharp, D., Braithwaite, N. St.: Looking into a plasma loudspeaker. IEEE Trans. Plasma Sci. **39**(11), 2146–21487 (2011)
17. Law, V.J., Chebbi, A., O'Neill, F.T., Dowling, D.P.: Electro-Acoustic Resonances within the kINPen-MED Atmospheric Pressure Plasma Jet. Presented at the 6th Chaos 2013 internal conference, Istanbul, Turkey, June 2013. In Press: Chaotic modeling and simulation journal, Jan 2014
18. Law, V.J., Nwankire, C.E., Dowling, D.P., Daniels, S.: Acoustic emission within an atmospheric discharge jet. In: Skiadas, C.H., Dimotikalis, I., Skiadas, C. (eds.) Chaos theory: Modelling simulation and applications. World Scientific Publishing Co Pte Ltd, Singapore (2011). ISBN: 9814350338
19. Levine, H., Schwinger, J.: On the radiation of sound from an unflanged circular pipe. Phys. Rev. **73**(4) 383–406 (1948)

Altruism and Identity

Burton Voorhees

Abstract This paper reviews some of the major approaches to the evolution of altruism and cooperation, addressing the question of how fitness reducing behaviors could evolve and become fixed in a population. An often overlooked important point is that human psychology must be taken into account, in particular the way that personal identity is tied to adherence to social and cultural norms and approved behaviors resulting in an identification of personal biological survival with a social group and the markers of that group.

Keywords Evolution of altruism · Cooperation · Group selection · Personal identity

1 Introduction

Altruism and morality have been issues of intense interest and contention at least since early Greek philosophers debated the question of virtue. What duty do we owe to our fellow human beings? Is there a universal morality, or are all moral and ethical codes culturally specific? How and where did moral sentiments originate, and how do they conform social behavior?

In modern evolutionary theory such questions are problematic. If genes are selfish and behavior is directed solely toward reproduction of genetic traits, how is it that altruism could have evolved? And without altruism, how could morality and cooperation arise, other than as forms of mutual self-interest? Yet altruism is sometimes observed in nature (e.g., [1–3]) and, cynicism to the contrary, humans seem to be altruistic in the extreme. We are, as indicated in the titles of recent books, "super cooperators" [4], "a cooperative species" [5]. How such behavior could evolve, in

B. Voorhees (✉)
Center for Science, Athabasca University, 1 University Drive, Athabasca, AB T9S 3A3, Canada
e-mail: burt@athabascau.ca

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_33,

spite of the loss of fitness implied by acting for the benefit of others at a cost to oneself, is a major subject of research.

Evolutionary studies of altruism and cooperation have an economic focus, offering explanation in terms of transmitting closely related genes (kin selection), selection enhancing behaviors related to expectations of reciprocity in one form or another (reciprocal altruism, mutualism), or behaviors sustaining group cohesion (group, or multi-level selection). Altruism may be individually costly, but it arises and becomes fixed in a population because, overall, it provides benefits to the social group outweighing individual costs. Individuals "pay" for the benefits of group membership with a reduction of inclusive fitness. But human altruism and cooperation seems to go beyond actions that provide biological or economic reward. Humans gain psychological benefits from extending help to others, and feelings of satisfaction at "doing a good deed" may provide sufficiently rewarding. "People cooperate not only for self-interested reasons but also because they are genuinely concerned about the well-being of others, try to uphold social norms, and value behaving ethically for its own sake. … Contributing to the success of a joint project for the benefit of one's group, even at a personal cost, evokes feelings of satisfaction, pride, even elation. Failing to do so is often a source of shame or guilt" [5, p. 1]. Nevertheless, the question remains: how could a costly, fitness reducing behavior evolve, much less flourished?

This chapter reviews some of the approaches taken to answer the question of altruism and links these answers to psychological theorizing about the nature of human self-consciousness. Briefly, there is a question of identity involved that has been noted (e.g., [5]) but not sufficiently emphasized in previous evolutionary literature. Humans congregate in marked groups (e.g., [6]) and will make personal sacrifices for the benefit of the group, even if it is not composed of close genetic kin and even if no personal benefit, either direct or indirect, accrues. This is explained, in part, by the human capacity for (and vulnerability to) identification with ideas, beliefs, symbolic cues, and other group markings. These group markings become a part of personal identity and, as such, evoke biologically based survival related behaviors.

2 Varieties of Altruism

The term altruism is used in different ways, only one of which refers to pure altruism; that is, costly behavior that bringing no compensatory return. Whether pure altruism exists or not is a matter of dispute (e.g., [7]). Other behaviors labeled altruistic are mutually beneficial cooperation (mutualism), kinship altruism, and reciprocal altruism. In each of these cases there is an assumed quid pro quo, compensation for the loss of inclusive fitness: acting altruistically towards close kin provides an enhanced possibility of genetic continuation; cooperation implies an expectation of benefits arising from a cooperative venture; and reciprocal altruism implies an expectation of favors returned for favors given. To this we can add a uniquely human form,

sentimental altruism, in which the payment for altruistic behavior is the experience of desired emotions and feelings that boost self-esteem [8].

A major attempt to explaining the evolution of cooperative and altruistic behavior is presented by Bowles and Gintis [5], who posit the following overall scenario.

1. Altruistic impulses are grounded in human social preferences (both other and self-regarding) and these preferences are supported and reinforced by associated social emotions, in particular guilt and shame.
2. Because group living is an essential condition for human survival, the biological source of social preferences and social emotions is natural selection acting on individuals living in groups.
3. Human groups differ in evolutionary success and this difference correlates with the degree of cooperation and altruistic behavior within the group.

They suggest, "The fact that more cooperative groups tended to survive and expand explains two key facts. The first is that altruistic individuals among our ancestors enjoyed enhanced reproductive success, resulting in the spread of altruism as a distinctive human trait. The second is that our altruistic dispositions motivate us to care about and help not only close family but even those related to us only distantly or not at all, as long as we share common group membership" [5, p. 51].

In order to support these claims, simulation and agent based models are used to test various game theoretic assumptions for the emergence of evolutionarily stable equilibriums supporting cooperative and altruistic behavior. They propose that such equilibriums are stabilized by the correlating mechanism of a system of social norms; and that an essential aspect of the social emotions supporting altruistic behavior is the desire to punish norm violators, and even to punish those who do not punish.

Since social norms vary between different groups, competition for resources between hunter-gatherer groups in a Paleolithic environment acts as a selective process at the group level. Groups in which social norms support within group cooperative behavior will tend to win such competitions. Within group adherence to cooperative norms, along with punishment of violators, produces a selective pressure on individual group members, leading to an increase in the frequency of altruistic individuals. Thus, a social niche is constructed favoring the evolution of altruistic and cooperative behavior. The success of any such process, however, relies on inter-generational transmission of social norms, which requires cooperative action.

An extended discussion of cultural transmission (cultural learning) is given by Sterelny [9], who argues that "hominins developed a new form of ecological interaction with their environment, cooperative foraging, and this ecological revolution led to positive feedback between ecological cooperation, cultural learning, and environmental change" [pp. 3–4]. Most research on the evolution of cooperation assumes the "Machiavellian hypothesis": that the major issue with cooperative action is to avoid free riders. For Sterelny, however, the major issue is coordination: while most models of and experiments on cooperation "explore the consequences of different patterns of interaction, the effects of punishment, or error, of group structure, of the effects of manipulation of rewards and costs. *They do not explore the mechanisms that generate the rewards of cooperation*. …[The] Machiavellian hypotheses thus

focus on this cognitive challenge of managing cooperation in an environment in which defection is a threat rather than the problem of coordinating, of organizing collective action so that it generates a cooperative profit" [p. 7].

The Machiavellian view is that a major evolutionary pressure driving human cognitive development was detection of cheaters. Sterelny points out that in a hunter-gatherer society, people will generally be aware of who the defectors are and the real issue will be coordinating cooperative efforts. Once cooperation is instantiated as a stable group behavior, the need to resolve problems surrounding coordination, rather than detecting cheaters, provide the primary selective pressure for further cognitive evolution: controlling (rather than detecting) cheaters is a problem of coordination.

Cooperative hunting is an example. Killing large animals with sharp sticks isn't easy—it requires substantial cognitive capacity: the ability to react as a group to contingent events, skills in hunting and tracking, the knowledge to manufacture weapons, as well as a good foundation of knowledge about animal behaviors and environmental conditions. Successful collective foraging requires far more knowledge than could be learned anew in each generation. In other words, social learning is a prerequisite if cooperative efforts are to provide benefits.

3 Social Learning and Cultural Transmission

Social learning is part of a positive feedback interaction: the cost of social learning is an expensive, extended childhood and adolescence in which cultural and survival related information is acquired. The benefit is that the information acquired allows adults to forage successfully, producing the surplus resources necessary to support the extended childhood learning period.

Sterelny proposes an apprentice model for social learning that could have arisen in early human evolution. Children associate with adults while they are carrying out their activities, and learn by imitation and trial-and-error, with occasional explicit instruction. He posits that this method of social transmission continued for an extended period until, about 50,000 years ago, it stabilized sufficiently that not only could information be transmitted from generation to generation, but this information could be modified and improved, and these improvements could also be transmitted. Thus in Sterelny's view, the symbolic revolution that began about 50,000 years ago is an indication of crossing a threshold. While many of the elements of this revolution can be identified earlier in human evolution [10], he attributes the onset of the symbolic revolution to establishment of not only the ability to reliably transmit cultural information, but also *the onset of the reliable transmission of cultural innovation*.

Lewis-Williams [10, pp. 95–96], on the other hand, associates this transition to the appearance of social distinctions between different individuals: "It was not cooperation but social competition and tension that triggered an ever widening spiral of social, political, and technological change." But ethnographic data shows that hunter-gatherer groups are egalitarian, with most competition arising between groups rather than within a group (e.g., [5]). Indeed, if cooperation is to arise and stabilize

in a population, within-group competition needs to be minimized and this social leveling is part of the function of the norms transmitted through social learning.

Read [11] refers to the Symbolic Revolution as an "organizational rubicon" in which new structures of social organization arose drawing on linguistic and cultural forms made possible through previously evolved cognitive capacities. Read associates this with the rise of cultural kinship systems. Modern ethnographic data on hunter-gatherer societies indicates that they are composed of relatively small residence groups within which social norms, encoded in culturally constructed kinship systems, maintain an egalitarian ethic. On available data, this is the sort of social organization that existed in the Paleolithic, providing a social niche for the evolution of cooperative and altruistic behavior.

Ethnographic evidence also indicates that modern hunter-gatherer societies have worked out effective means to control cheats, free-riders, and bullies. Methods of control include gossip, ridicule, ostracism, physical punishment, and, on occasion, murder. Lee [12, p. 246] quotes a !Kung san informant "When a young man kills much meat, he comes to think of himself as a chief or a big man, and he thinks of the rest of us as his servants or inferiors. We can't accept this. We refuse one who boasts for someday his pride will make him kill somebody. So we always speak of his meat as worthless. In this way we cool his heart and make him gentle."

Willingness to punish defectors, however, faces a commitment problem. An agent at a given time commits to punish a defector at a future time, even though actually acting on this commitment will be costly in terms of inclusive fitness. But if the threat of punishment is to be credible, there must be a credible guarantee that defection will be punished. Likewise, cooperation in joint endeavors requires commitment to keep agreements and not defect should the temptation arise to do so.

Commitment can be reinforced externally by being made public. One way of doing this is through body markings such as scarification, tattoos, or similar group markings. These indicate commitment to the group and, if they are permanent marks, greatly increase the danger associated with ostracism. Another external commitment mechanism is boasting. If an agent reneges after publically committing to an action, they expose themselves to ridicule and loss of reputation.

Internal psychological commitment mechanisms are found in the social emotions, in particular the powerful emotions triggered by violations of trust or fairness. Experiencing these emotions sets internal controls on commitment. Further, since people tend to be good at recognizing capacity, poor at recognizing dispositions, reliable signal of trustworthiness are important. Because the external displays of emotions related to trust and fairness are hard to counterfeit, these can provide signals of trustworthiness, of willingness to cooperate in group enterprises, and to punish defectors.

Another powerful commitment building mechanism, Sterelny suggests, is hunting. Hunting, in a Paleolithic environment, is high risk and high arousal. Participation in hunting, which is essentially universal among males in hunter-gatherer bands, builds bonds of trust among participants. In this case, however, it is a gradual thing: "The mutual experience of hunting and sharing does not signal *a preexisting bond* of loyalty and attachment that makes trust and trustworthiness likely. Rather, it incrementally

creates those psychological bonds. Trust and trustworthiness are both the products of costly, high-arousal activity" [p. 116].

High cost of an activity increases arousal, which increases the psychological effect of the activity on participants. Hunting, group song and dance, and initiation rites not only *signal* commitment, they have lasting cognitive and affective effects. Trust is built incrementally and, over time, the bonds of commitment become economically powerful. Alliances are built that can be counted on because the participants have made substantial investments. This changes the payoff matrix when faced with a temptation to defect. Defection may bring temporary advantage, but the cost is loss of the trust network. Given the role of reputation in a linguistic species that gossips, this is a powerful constraint on defection. Trust is constructed through social interactions in cooperative ventures through which others come to see an agent as trustworthy. The effort involved in building a good reputation provides a disincentive to defect since this investment pays off only so long as the person maintains their trustworthy reputation. A good reputation is not something to throw lightly away.

A major change in human evolution occurred with the emergence of culturally self-identified groups. Individuals moved from simply living within a group to identifying themselves as members of that group and adopting group markings signifying this membership. This required new cognitive and social skills, providing the basis for cultural kinship, and [11, p. 14] "Cultural kinship became the innovative framework through which new systems of social organization are defined and expressed."

Cultural kinship relations carry expectations of behavior, they impose pre-established commitments, characterized by Fortes by an axiom of amity: "kin are expected to be mutually supportive even if only by virtue of being recognized as kin through a conceptual system of kin relations, independent of biological kin distance" (from [11], p. 15). A kinship relation provides an expectation of trust—a defector who violates kinship obligations faces strong social sanctions.

Kinship also provides a template for coordinated activities "Like the species boundary that relates to the maintenance of coordination among genetically based behaviors, the cultural boundary relates to the maintenance of coordination among individuals for culturally based behaviors. Individuals enculturated within the same cultural milieu understand the meaning of the behaviors of others in a similar way, and others respond to an individual's behavior in a similar manner through shared, cultural meaning of actions" [11, p. 68]. In other words, coordination of activities within the cultural group is simplified, behaviors can be anticipated, and if anticipations are not met, the offender can be sanctioned.

One aspect of kinship expectations and obligations relates to the sharing of food. Food sharing, especially by successful hunters, is an activity that reduces within group selection and there are set cultural rules and norms for sharing of food and other collective resources. "The sharing of hunted animals is not based on an individual decision by the hunter whether he should share, but takes place in a culturally specified manner. This takes decisions about sharing out of the hands, as it were, of the hunter(s) and thereby transforms what otherwise may be behavior that is unpredictable and/or subject to substantial variation, depending on individual interests in, or proclivity towards sharing, into predictable outcomes when hunting has been successful"

[11, p. 86]. Hunting is a high-risk endeavor. The cultural rules for sharing the results of a successful hunt effectively average the risk over all hunters, regardless of their individual skill or luck. It also reduces the potential for intra-group conflict since everybody knows before hand what share of the meat they are justified in receiving. Meat sharing also promotes an egalitarian social structure. Successful hunters cannot monopolize meat in order to gain hierarchical status.

Read uses the Netsilik Inuit as an example of how cultural systems such as kinship and, in the Netsilik case, the institution of "sealing partners" for hunting seals, constrain behaviors through prescribed forms of meat sharing: "A Netsilik hunter does not decide whether it is in his interest to share a seal he has killed, but shares it in the prescribed manner since to do otherwise contradicts his understanding of what it means to be a sealing partner. …To act otherwise is equivalent to him saying that he (and his family) are opting out of the system of sealing partners" [11, p. 90]. This appears to be a form of reciprocal altruism since a hunter who has killed a seal and shares the meat as required can expect the same sharing when another of his sealing partners makes a kill. But there is a difference—a hunter who defects from the sealing partner system will face social consequences for violation of cultural rules and the response will not be tit-for-tat (i.e., other hunters refusing to share meat with him) since that would constitute a further violation of cultural norms. Rather there will be a collective punishment that, if it extends to ostracism, could be lethal.

The essential point is that defection is a violation of a system of cultural norms that is an embodied aspect of each individual's social identity. A defector is not only acting selfishly, *he is contradicting his identity as a member of the group*.

Culture, from Read's perspective, is the existence of a set of understood concepts that frame behavior. Culture transmits norms, but these are not behavioral maxims, they are specifications of who to cooperate with and what sort of expectations can be had of others. These norms are not simply *acquired* through imitation; they are *imposed* by a process of enculturation, which occurs simply in virtue of being born into and coming to adulthood within a given culture.

Social interactions based on culturally defined roles, kinship or otherwise, depend on all individuals sharing the same cultural background, hence sharing the same concepts about proper behavior. In at least some hunter-gatherer groups there is little evidence that altruistic behavior extends beyond the cultural group. Referring to the !Kung san, Weissner [13] notes that "There is little evidence… that it is a part of human psychology [for them] to be willing to engage in altruism… in a social and cultural vacuum. When the faces and forces of culturally defined institutions are reintroduced, sharing and giving resume." In other words, altruistic behavior is, in many cases, carried out simply because it is an expected behavior according to some system of cultural ideals relating to how one behaves towards other group members, and these ideals have been introjected as an aspect of personal identity.

Because of this, once cooperative and/or altruistic behavior becomes fixed in a cultural kinship system, there is no longer a need to explain it in terms of biological selection acting on individuals. Rather, it is only necessary to examine the way in which other cultural ideas regarding cooperative behavior are related to the primary kinship system. Failure of members of a cultural group to follow prescribed

cooperative or altruistic behaviors is no longer simply a game theoretic calculation of individual interest; it is a social and psychological opting out of the cultural matrix itself, a matrix that is intrinsic to the individual's identity.

This means that it is necessary to make a clear distinction between intrinsically group-level properties, and group-level properties that emerge from individual behavior. The latter are still subject to individual level biological selection, even though the emergent group property alters the environment within which this selection takes place. Group-level cultural idea systems, on the other hand, are intrinsic group level properties that are subject to group-level selection, leading to evolution at the group level. These group-level traits don't arise from individual traits, rather *they provide fitness benefits for individuals only in virtue of individual collective conformity to the practices and norms instantiated within the system*. It is not that genetic selection favors cooperation giving rise to cooperative group norms, rather cooperative group norms are embodied by individual group members as an aspect of their identity as members of the group, and this identity is directly connected to their continued survival as an individual.

While group-level traits are subject to group level selection, they also form part of the environment that an individual group member experiences, resulting in top down selective pressures on individual fitness. This selection favors conformity to group norms, possibly instantiated through evolutionary refinement of the social emotions. This dual level process explains why, if selection is for in-group cooperation, traits favoring altruistic cooperation have not gone to fixation resulting in a eusocial species such as the naked mole-rats [14]. Top down constraints are only partially effective, and the environment provided by group-level traits also selects for successful cheaters, and in response, for detection of cheaters [15]. But the cheater is at a disadvantage in this sort of arms race because successful deception requires greater cognitive resources than does honest cooperation.

4 How Are Social Emotions and Cultural Systems Introjected?

Bowles and Gintis [5] suggest that social emotions such as shame and guilt reduce the necessity of costly punishment, through their connection to contempt. To feel contempt from others is a direct threat to self-image, and a public expression of contempt tarnishes reputation. Defectors not only exclude themselves from their peer group, they are engaging in actions that, to the extent they identify as a member of the group, go against their self-image, and this evokes negative emotional responses.

In this regard, Bowles and Gintis raise two basic questions: (1) since social emotions are often altruistic, they would be expected to disappear in "any dynamic in which the higher payoff trait tends to increase in frequency," and (2) "How could it ever be evolutionarily advantageous to bypass one's cognitive decision making capacities and let behavior be influenced by the visceral reactions associated with one's emotions?" [5, p. 191]. In regard to the second question, they suggest that immediate, emotion-driven responses save cognitive effort under conditions of

limited information and limited decision time. As for the evolution of social emotions, they suggest that emotions such as guilt and shame—because they reduce the cognitive load in decision making by providing an immediate corrective for behavior now that would have negative consequences in the future—provide an adaptive advantage even though they also promote the fitness reducing aspect of altruistic behavior.

One question remains. What is the vehicle across which the social emotions manifest and how do they gain the intense motivational power they wield? For purposes of discussion, I will draw substantially from Damasio's [16] hypothesis on the emergence of the self. In this framework the self is a process rather than a thing, and there is a distinction between the self-as-object ("the dynamic object constituted by certain workings of minds, certain traits of behavior, and a certain life history") and the self-as-knower ("the process that gives a focus to our experiences and eventually lets us reflect on those experiences"). The self-as-object, quoting William James, is thought of as "the sum total of all that a man could call his—'not only his body and his psychic powers, but his clothes and his wife and children, his ancestors and friends, his reputation and works, his lands and horses, and yacht and bank account"'. Going beyond this, "what allows the mind to know that such dominions exist and belong to their mental owners—body, mind, past and present, and all the rest—is that the perception of any of these items generates emotions and feelings, and, in turn the feelings accomplish the separation between the contents that belong to the self and those that do not" [p. 9].

The basic issue Damasio attempts to deal with is how a self-process arises in human brains. He proposes a framework based on a series of assumptions:

1. The conscious mind is intrinsically based in the body and the body image in the mind forms a *protoself*.
2. The structures and processes of the protoself are the result of information communicated to the brain from the body, and the return impulses sent from the brain to the body.
3. The first, most elementary aspects of the protoself are *primordial feelings*, which arise continuously and spontaneously. These provide the direct experience of a living body. They reflect feelings such as pain and pleasure and are the basis of all feelings and emotions. The primordial feelings are the first manifestations of sentience.
4. The self is built up in a series of steps grounded in the protoself. When any relation appears between the organism and an object, this modifies the protoself, including the primordial feelings, and the unfolding sequence of images in the mind describing this relation and its effect is an element, or pulse, of the *core self*.
5. Following on this, there is biographical knowledge which gives rise to the *autobiographical self*, generated from "complexes of images that generate pulses of core self."
6. The fundamental underlying principle is homeostasis: maintaining and sustaining life within viable limits. This homeostatic imperative exfoliates into

individual life, society, and culture, giving rise to regulatory devices oriented to the sustainment and preservation of "biological value."

The main function of the brain, for Damasio, is map-making, an essential tool of life management. Maps are made through interactions with the environment, through cycles of feedback and feed forward between perceptual and action. The mind itself is a result of this map-making process. The sequences of images flowing through the mind gain salience according to their value for the individual. That value comes from the set of dispositions that orient homeostatic regulation, and from valuations that have been acquired through life experiences, encoded in mental maps.

The essential map, from the perspective of an emergent self-process, is the body image, which is altered by events that influence the body, and which acts back on the body psychosomatically. Through this image the brain is able to know body states and responsively control bodily reactions in order to maintain essential physiological and other parameters within their desirable homeostatic range. The fundamental vehicles for this process are the primordial feelings.

The brain is also able to simulate body states, both our own and those of others, providing a basis for empathy. There are feedback/feed forward loops connecting body states, actual movements, somatosensory representations of movements, visual representations of movements, and memory. "The living body is the central locus. Life regulation is the need and the motivation. Brain mapping is the enabler, the engine that transforms plain life regulation into minded regulation and, eventually, into consciously minded regulation" [p. 107].

Feelings and emotions, however, are distinct. "Emotions are complex, largely automatic programs of *actions* concocted by evolution. These are complemented by a *cognitive* program that includes certain ideas and modes of cognition, but the world of emotion is largely one of actions carried out in our bodies, from facial expressions and postures to changes in the viscera and internal milieu. Feelings of emotion, on the other hand, are composite *perceptions* of what happens in our body and mind when we are emoting" [p. 109]. Feelings are the mental images of viscera-autonomic, kinesthetic, and other bodily responses that have an evolutionary basis. "Feelings of emotion are composite perceptions of (1) a particular state of the body, during actual or simulated emotion, and (2) a state of altered cognitive resources and a deployment of certain mental scripts. In our minds, these perceptions are connected to the object that caused them" [p. 116]. All feelings of emotions are variations on the primordial feelings, augmented by culturally overlaid cognitive beliefs and scripts.

This yields a system in which the bottom up sense of self-as-body, arising from the primordial feelings, gains identity through unique experience (personal identity) and top down cultural constraints (social identity). Because the resulting self is the result of a process grounded in the primordial feelings; feelings that are genetically evolved to promote survival, the integrity of this self-identity will be protected with an intensity rooted in basic survival instincts. This self is constructed, in part, from cultural inputs, and these inputs set perceptual and cognitive cues for evocation of social emotions. Thus, social emotions gain their motivating power from the survival related feelings that their evocation arouses. The perception that a group member

is violating group norms, for example, yields a cognitive script that cues emotional responses, expressed in the body (and thence, back in the mind) as feelings motivating reactions of indignation, condemnation, and punishment.

References

1. de Waal, F.: Good natured: the origin of right and wrong in humans and other animals. Harvard University Press, Cambridge (1986)
2. Heinrich, B.: Ravens in winter. Summit Books, New York (1989)
3. Ratniaks, F.L.W., Helanterä, H.: The evolution of extreme altruism and inequality in insect societies. Philos. Trans. R. Soc. B **364**, 3169–3179 (2009)
4. Nowak, M.A., Highfield, R.: Supercooperators. Free Press, New York (2011)
5. Bowles, S., Gintis, H.: A cooperative species. Princeton University Press, Princeton (2011)
6. Richerson, P.J., Boyd, R.: Not by genes alone: how culture transformed human evolution. University of Chicago Press, Chicago (2005)
7. Lichtenberg, J.: About altruism. Philos. Public Policy Q. **28**, 2–6 (2008)
8. Margolis, H.: Selfishness, altruism and rationality. Cambridge University Press, Cambridge (1982)
9. Sterelny, K.: The evolved apprentice. MIT Press, Cambridge (2012)
10. Lewis-Williams, D.: The mind in the cave. Thames and Hudson, London (2002)
11. Read, D.W.: How culture makes us human. Left Coast Press, Walnut Creek (2012)
12. Lee, R.B.: The Kung san: men women and work in a foraging society, p. 246. Cambridge University Press, Cambridge (1979)
13. Wiessner, P.: Experimental games and games of life among the Kalahari Bushmen. Curr. Anthropol. **50**(1), 133–138 (2009)
14. Honeycutt, R.L.: Naked mole-rats. Am. Sci. **80**(1), 43–53 (1992)
15. McNally, L., Jackson, A.L.: Cooperation creates selection for tactical deception. Proc. R. Soc. B **280**, 20130699 (2013)
16. Damasio, A.: Self comes to mind. Pantheon, New York (2010)

Synchronization of Circadian Rhythms at Scale of Gene, Cell and Whole Organism

Andrey Zakharov and Dmitry Bratsun

Abstract Three characteristic scales of a biological system are distinguished in the chapter: microscopic (gene's size), mesoscopic (cell's size) and macroscopic (organism's size). For each case the approach to modeling of the circadian rhythms is discussed on the base of a time-delay model. The stochastic description has been used at the gene's scale. The deterministic description within the spatially extended model has been suggested on the mesoscopic scale. Macroscopic effects have been analyzed within the discrete model describing the collective behaviour of large amount of cells. The effect of collective rhythms synchronization for each case has been studied. The problem of cross-linking of the results obtained at different scales is discussed.

Keywords Synchronization · Circadian rhythms · Time-delay · Intrinsic noise · Individual based models · Reaction-diffusion systems

1 Introduction

Biological rhythms are periodically repeated changes of biological processes, which are quite characteristic for living matter on every level of its hierarchy starting from molecular and subcellular up to biosphere on the whole. Rhythms development in a living organism is closely connected with its adaptation processes to the environment during the evolution. Now it becomes clear that these rhythms are embedded in the genetic structure. So if the external effects are eliminated, the periods of these rhythms differ from the periods of the corresponding rhythms of the environment [1].

A. Zakharov (✉) · D. Bratsun
Theoretical Physics Department, Perm State Pedagogical University, Perm, Russia
e-mail: az1211@mail.ru

D. Bratsun
e-mail: dmitribratsun@rambler.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_34,

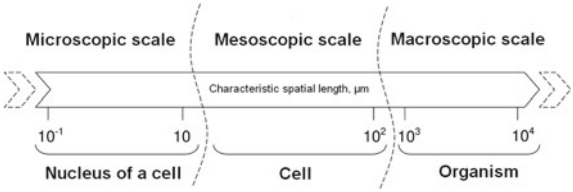


Fig. 1 Characteristic spatial scales of a biological system

Although rhythms classifications vary, in this article we concentrate on the rhythms, which synchronize with the daily changes of the environment and are called circadian.

At the time, when the genetic nature of the circadian rhythms was unclear, these rhythms were assigned to a particular scale of organism's organisation. For example, a monograph [2] assigns the ultradian rhythms to cells and tissues. Circadian rhythms are developed at the scale of the whole organism. According to this hypothesis, these circadian rhythms are drivers-rhythms: they are labile to external factors influence, can synchronize with them and affect the subordinate driven rhythms. However, the identification of the genetic mechanism of circadian oscillations [3] made scientists realise that understanding was wrong. Circadian rhythms mechanism works even at the scale of one or several genes by revealing itself in RNA and protein fluctuations in transcription/translation processes. Finally, as soon as transport protein gets through the cell membranes and starts its cellular interaction, circadian oscillations can inevitably be developed at the intercellular scale. At the organs' scale the signals from separate cells should be synchronized, thus developing unified rhythms for the whole organism. Notice that the issue of spatial synchronization of a large amount of collaborating oscillators is quite popular among the physicists [4], but one could hardly meet the discussion of this issue among the biologists. Major part of literature on circadian rhythms is concentrated on the temporal organisation of rhythms.

In this chapter we discuss the approaches to circadian rhythms modeling at different spatial scales of a living organism, and study the forms of rhythms synchronization, which are developed at different scales of system functioning.

2 Characteristic Scales of a Biological System

Let us point out three main biological system description scales, which are based on the characteristic scales of structural units of the whole organism (Fig. 1). Aside from the big cores of the reproductive cells with their gigantic sizes (up till 500 μm) cell core, which stores genetic information, is known to be of 1 till 10 μm in size at eukaryote. The transcription process, which presupposes data reading from the particular cells by RNA polymerase and synthesis of mRNA, occurs in a core. Then the molecules of mRNA leave the core and get into ribosome, where the proteins are synthesized. Ribosome is twice as small as a core—about 0.03 μm. The proteins may return to the core and influence the transcription process by interacting with

genes promoters. Therefore, here the characteristic scale of the spatial processes is determined by the core's sizes. However, the number of molecules interacting during the genetic processes of transcription/translation is far from being big—for example, the speed of RNA polymerase's machinery is only 50 nucleotides per second. So even minor fluctuations in mRNA and protein concentrations could have significant influence on the general dynamics of the system [5, 6]. Thus, system description at the scale of one or several genes should principally be stochastic. At the same time here one could not take into account the spatially extended dynamics of the molecular cloud, since only a small number of elements (tens or hundreds of molecules) are concerned. This description scale shall be termed *microscopic* (Fig. 1).

At the next scale of biological system, which shall be termed *mesoscopic*, a cell or several cells, which exchange the signals, are identified. A typical size of a eukaryote's cell is about 10–100 μm (Fig. 1). How many molecules of particular protein are there in a cell? For example, yeast has been carefully analysed in [7]. On average 4-digit or 5-digit numbers have been obtained for different types of protein. For example, the number of a gene *frq* (internal classification YDR373W), which is responsible for maintaining the circadian rhythms in yeast and in some other fungi, has been found to be 7160 per one cell [7]. It means that as far as the role of fluctuations is minor here, one may neglect the stochastic properties of the system. However, keeping in mind that the size of one protein molecule is about 0.001–0.01 μm we may introduce spatially extended model to describe the protein cloud dynamics. The diffusion coefficient for different protein monomers at their low concentration in water solution is known to be about 10^{-7} cm^2/s. This value is a bit lower than the standard value as the protein monomers are large heavy molecules consisting of hundreds and thousands of amino-acid residues. The protein diffusion in a cell cytoplasm is limited by the intracellular space structure. Now it is clear that the cytoskeleton of a cell consisting of actinic filaments is a complex viscoelastic medium with fluidization at certain conditions [8]. The diffusion coefficient value of protein in this intercellular frame is not known at the moment, but it is sure to be lower that the above mentioned value for water. Thus, we believe that at the mesoscopic scale the spatially extended deterministic reaction-diffusion model is the most appropriate to describe the system.

Finally, *macroscopic* scale presupposes the description of the ensemble of a large number of interacting cells (Fig. 1). This is the scale of an organ or even a whole organism. And here the system becomes discrete again since a cell is an elementary unit of organism structure, which can independently exist, grow and be reproduced. The characteristic value of chemical signals diffusion, which occurs between the cells, is quite small due to the influence of cellular membranes. It means that the exchange between the cells is rather slow and it is quite possible to apply the discrete model. This approach enables us to stand away from the mesoscopic dynamics inside the cells and to count only the intercellular differences.

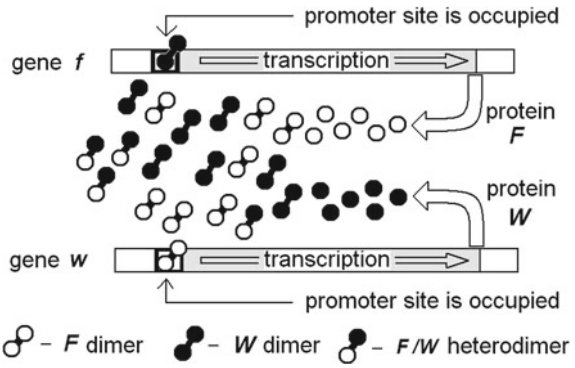


Fig. 2 Scheme of protein interaction in circadian oscillations

3 Stochastic Simulations at the Microscopic Scale

We shall use the dynamic model of circadian rhythms proposed in our previous chapter [9]. The model is rather general by nature, although it was originally suggested to describe the circadian rhythms of the organism *Neurospora crassa* [10]. For example, a model of the similar type has been developed for the circadian oscillations of a fly *Drosophila* [11]. The time-delay effect of protein synthesis reactions in the transcription and translation processes of genes is the key element of oscillations mechanism (Fig. 2). These processes are both very slow and consist of multistage biochemical reactions involving the sequential assembly of long molecules. Thus, these processes are long in time and particularly time-delayed. It is evident that the delay prevents the system from achieving equilibrium, and results instead in the familiar limit cycle oscillations. The deterministic and stochastic properties of gene regulation taking into account the non-Markovian character of gene processes were studied in [12].

The simplified graphical depiction of a genetic mechanism of circadian rhythms is presented in Fig. 2. This pair of the principal genes responsible rhythms has been identified for some organisms. For example, for *Neurospora* these genes are *frq* and *wcc*, for *Drosophila—per* and *dclock* [11]. Generally, the circadian rhythms are characterized by a number of properties (autonomy, temperature compensation, etc.) determined by tens of other genes. However, the specified pairs are the principal ones for rhythms maintenance. Table 1 gives a complete list of biochemical reactions. The model has been described in more detail in [10].

As it was mentioned earlier, at the microscopic scale (1–10 μm, Fig. 1) we may ignore the spatial effects and consider only stochastic properties of the system. As it is known, in stochastic researches of gene processes two types of noise are pointed out: extrinsic noise, which is generated outside a cell and connected with the intercellular differences, and intrinsic noise, connected with stochastic nature of the ongoing chemical reactions, temperatures fluctuations, etc. It is clear that at the scale of a gene we should take into account the effect of the intrinsic noise.

Table 1 A list of genes transcription reactions. Here $k, k_1, k_{-1}, k_2, k_{-2}, B_F, B_W$ are the speeds of the corresponding reactions

1	Dimerisation	$F + F \xrightarrow{k_1^F} F_2 \quad , \quad W + W \xrightarrow{k_1^W} W_2$
2	Dedimerisation	$F_2 \xrightarrow{k_{-1}^F} F + F \quad , \quad W_2 \xrightarrow{k_{-1}^W} W + W$
3	Promoter binding	$D_0^F + W_2 \xrightarrow{k_2^F} D_1^F, \quad D_0^W + F_2 \xrightarrow{k_2^W} D_1^W$
4	Promoter unbinding	$D_1^F \xrightarrow{k_{-2}^F} D_0^F + W_2, \quad D_1^W \xrightarrow{k_{-2}^W} D_0^W + F_2$
5	mRNA synthesis	$D_1^F(t) \xrightarrow{k_F} D_1^F + F^{t+\tau}, \quad D_1^W(t) \xrightarrow{k_W} D_1^W + W^{t+\tau}$
6	Protein degradation	$F \xrightarrow{B_F} \varnothing, \quad W \xrightarrow{B_W} \varnothing$
7	Non-linear degradation	$F + W \xrightarrow{k} \varnothing$

Table 2 Model parameters

τ	k	k_F	k_W	K_1^F	K_2^F	K_1^W	K_2^W	B_F	B_W
6 h	30 $\text{nM}^{-1}\text{h}^{-1}$	8 nM/h	4 nM/h	5 nM^{-1}	5 nM^{-1}	5 nM^{-1}	5 nM^{-1}	0.3 h^{-1}	0.4 h^{-1}

Gillespie's algorithm [13], which is a variety of Monte-Carlo's methods, is a convenient tool for stochastic genetic research. This method is especially actively applied, where a relatively small amount of molecules is involved. The numerical solution received by Gillespie's method is known to statistically reproduce the exact solution of the master-equation. It should be noted here that the traditional version of Gillespie's algorithm was developed only for Markovian systems. The algorithm was modified for the time-delayed systems in [12].

Let us present the results of stochastic calculations based on modified Gillespie's algorithm and parameter values listed in the Table 2. Figure 3a shows time-based diagrams for the total number of monomers of both types, which have been obtained at numerical calculations of a deterministic dynamic system (see below). Figure 3b illustrates the results of the stochastic simulations for the same parameters values as it is in a deterministic case. It is clearly seen that the basic mechanism of the oscillations works very well, although there are a small number of the molecules involved in the dynamics and significant fluctuations. The mechanism is based on the synchronization between variations of protein F and W being oscillated in anti-phase. Since productions of proteins are closely connected via positive feedback (see Fig. 2) it looks natural. System stochasticity is clearly seen in fluctuations, which achieve 20–40 % from the fluctuation magnitude in the deterministic case. Fourier spectra of signals are given in the *Insets*. It is seen that the stochastic signal has a stable maximum, which corresponds to the period of approximately 23 h. This proves the basic mechanism of circadian oscillations is robust with respect to even large fluctuations excited by intrinsic noise.

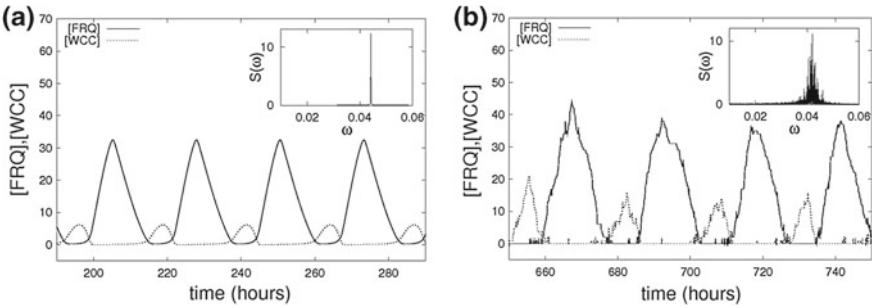


Fig. 3 **a** Time dependence of the total number of F monomers (*solid line*) and of W monomers (*dotted line*) deduced at the numerical integration of the deterministic model (1–2). **b** Time evolution of the corresponding stochastic system derived with the help of Gillespie's algorithm, which has been modified for non-Markovian processes. Fourier spectra of signals are given in *Insets*. The system parameters values are the same in both cases (Table 2)

4 Spatially Extended Model for Mesoscopic Scale

As it was mentioned above, modeling of biorhythms synchronization at the scale of one or several cells (10–100 μm, Fig. 1) presupposes the development of a spatially extended model. We show in [9, 10] that the set of biochemical reactions listed in the Table 1 can be reduced to the following two-variable reaction-diffusion system:

$$\frac{\partial F}{\partial t} = \frac{1}{1+4K_1^F F}\left(k_F \frac{K_1^W K_2^F W^2(t-\tau)}{1+K_1^W K_2^F W^2(t-\tau)} - B_F F - kFW\right) + D\left(\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2}\right), \quad (1)$$

$$\frac{\partial W}{\partial t} = \frac{1}{1+4K_1^W W}\left(k_W \frac{K_1^F K_2^W F^2(t-\tau)}{1+K_1^F K_2^W F^2(t-\tau)} - B_W W - kFW\right) + D\left(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 W}{\partial y^2}\right), \quad (2)$$

where D is the coefficient of protein diffusion inside the cell. It should be noted that the model (1–2) is not equal mathematically to a set of reactions given in the Table 1, since it has been based on the assumptions that some reactions are quick and some are slow [9]. Thus, the reagents involved into the quick reactions achieve the state of local statistic equilibrium very quickly against slowly changing values. So, one should use either the original system of kinetic chemical reactions (Table 1), or simplified deterministic model (1–2) depending on the system scale and a set of assumptions.

The initial-boundary value problem (1, 2) has been solved in the domain Σ: $(0 < x < 200, 0 < y < 200)$ by a finite difference method described in detail in [14]. The equations have been approximated on a rectangular uniform mesh 400×400 using a second order approximation for the spatial coordinates.

Since the proteins F and W are always in anti-phase, we can choose only one of them to illustrate the system dynamics. Figure 4 presents the typical wave

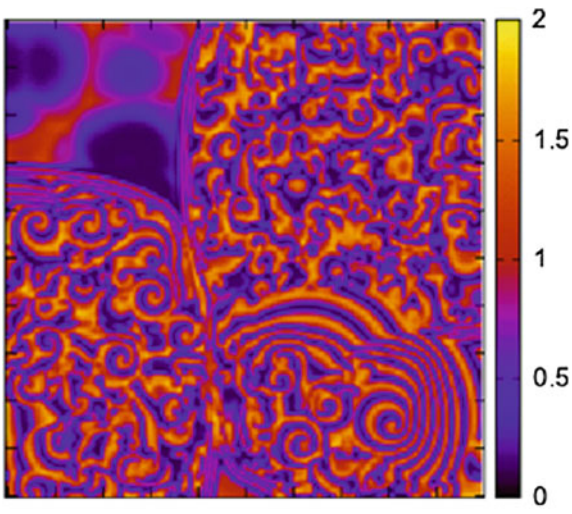


Fig. 4 Typical image of the wave structures developed within the intracellular space due to random initial conditions. $D = 0.01$

pattern formed by the concentration of F protein. The nonlinear dynamics of spatially extended system consists of two distinct oscillatory modes. One is the quasi-standing wave pattern visible in upper left corner of Fig. 4. The second oscillatory mode is a spiral traveling-wave pattern. These waves arise from selected initial disturbances and travel outward in all directions from its source. It continues until the spiral wave pattern occupies the entire domain Σ. Note that if the front of wave looks more or less orderly, by entering deeply inside the secondary instability area there have appeared numerous secondary centers of the excitation of the spiral waves. The nonlinear interaction between them leads to the formation of chaotic pattern (Fig. 4).

Many biological processes are carried out by signal transmission through cell membrane, which connect a cell with the environment. The membranes play the key role in metabolism by segregating non-organic ions and organic molecules. The changes in the levels of these flows are known to be one of the mechanisms of influence on the chemical composition and chemical structure inside a cell. This process has been modeled in the following way: in the domain Σ particular areas, which are conditionally called "cells" with the running reactions (1–2), are pointed out. The modeled cells are square and have a fixed size, because at the moment we are not interested in the mechanics of the cells themselves. The cells' border is not homogeneous: most of its part is impenetrable for a reagent, the other part has diffusion permeability through "membranes". For the sake of simplicity we have not introduced a separate type of protein responsible for the transportation of the circadian signal outside. In the suggested model the transportation function is performed by the same protein F. Its partner, protein W, may move only inside the cell space.

Now let us move to the results of the modeling. Figure 5 shows the situation, when only one cell generates the circadian rhythms (upper row). In other three cells the transcription mechanism does not work. As it was stated above (see Fig. 4), in this

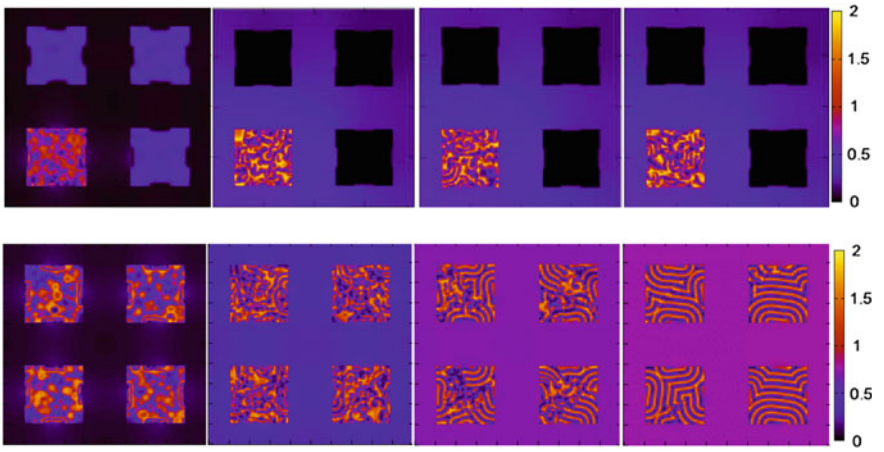


Fig. 5 Evolution of the concentration of the F protein. The frames from *left* to *right* and from *up* to *down* correspond to times $t = 200, 5000, 10000, 20000$ respectively. There are four "cells" in the domain of integration, but only one of them is switched on (*upper row*); The spatial synchronization of oscillations due to intercellular communications (*bottom row*)

case the system is in the state of the phase turbulence, which lasts as long as the external signal does not occur. The cell continues its functioning in its usual mode of chaotic spatial waves, although the protein F flows through the membranes of the working cell into the intercellular space. When all four cells (Fig. 5, bottom row) start functioning at the same conditions, then the system starts to change qualitatively. Spontaneous spatial synchronization has been proved to occur, when the protein concentration in intercellular space achieves a definite level of concentration $F = 0.8$, which equals to about an average value, with the concentration field in cells oscillating around it. It makes the intercellular diffusive space common for all cells.

5 Individual Based Model for Macroscopic Scale

The analysis of the biorhythms synchronization at the scale of an organ or even the whole organism presupposes the description of functioning and interaction of a large amount of cells. We have applied the chemo-mechanical model of epithelial spreading suggested in a recent chapter [15]. Epithelial tissue is a layer of cells covering the surface of an organ or body. Thus, one may use only quasi-two-dimensional system in modeling the epithelium behaviour, that makes the calculations easier.

The model described in the chapter [15] includes the calculation of separate cells dynamics, which are presented in the form of polygons (the system has been dimensioned in such a way, as hexagonal cell is the most probable form of a cell, although other forms of polygons are also possible). The cells are closely located to each other

forming solid two-dimensional epithelial surface. The model has a set of properties, which are suitable to simulate the behaviour of real epithelium:

- Possibility to change the cell's size in the process of tissue evolution (for example, wound healing) and to change the local mechanical properties of environment;
- Possibility for the total number of cells to spread in the system by their division in particular evolution conditions;
- Possibility for the cells to move by the mechanism of intercalation;
- The calculation of the dynamics of substances concentration, which participate in the regulation of tissues activities, for every cell of a community;
- Exchange of chemical signals done between the neighbouring epithelial cells through common border;
- Effect of cells polarisation, which occurs spontaneously or under the influence of the external conditions.

Thus, every cell in the model is under the effect of several chemo-mechanical influences, which make it develop together with the whole system. Fine structure of the spatially extended effects connected with the heterogeneity of the fields inside cells is not identified in the model, as all the fields inside every cell depend on time only. However, the model can reflect the pattern formation at the larger scales in comparison with a cell's size. This model can be classified as an individual based model demonstrating collective effects with individual behaviour of separate system elements. We have improved it with the calculation of the circadian rhythms in every cell of a community using the following equations:

$$\frac{dF_i}{dt} = \frac{1}{(1+4K_1^F F_i)}\left(k_F \frac{K_1^W K_2^F W^2(t-\tau)}{1+K_1^W K_2^F W^2(t-\tau)} - B_F F_i - k F_i W_i\right) + \alpha \sum_j L_{ij}(F_j - F_i), \quad (3)$$

$$\frac{dW_i}{dt} = \frac{1}{(1+4K_1^W W_i)}\left(k_W \frac{K_1^F K_2^W F^2(t-\tau)}{1+K_1^F K_2^W F^2(t-\tau)} - B_W W_i - k F_i W_i\right), \quad (4)$$

with $\alpha = 0.1$ being a coefficient of the protein F transfer through cellular membrane.

Term responsible for the protein transfer from a cell to a cell in the Eq. (3) is written in the form of a simple differential proportion with L_{ij} having the value of border length between i and j cells. Only cells neighboring i cell are summed up in the formula (3). If the level of protein F concentration in the given cell is higher than the one of its surroundings, then the flow of protein molecules goes outside. Otherwise the protein inflow goes from outside. After the division of any one cell into two parts new cells inherit the phase of the circadian rhythm of the parent cell.

Now let us move to discussing the results of modeling. Figure 6 shows several frames of evolution of epithelium consisting of 1600 cells originally. At the very beginning the phase of circadian rhythm has been set randomly in cells, while the external effects on the system have been excluded. Such an approach enables us to identify probable unique forms of the collective cells behaviour to synchronize the fluctuations. The numerical results show that in the case of a large amount of cells a complete synchronization, meaning total alignment of the oscillation phases

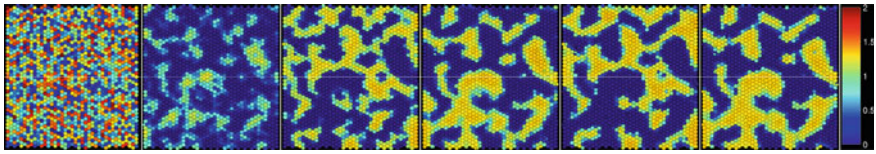


Fig. 6 Clustering of circadian rhythms: oscillations of the F protein in epithelial tissue consisting of more than 1600 cells. The random distribution of initial phases in cells has been applied as the initial condition. Time moment $t = 0, 25, 50, 75, 100, 125$ are presented consistently

in all cells cannot be achieved. Instead here the macroscopic effect of clustering is revealed—the cells develop two approximately equal communities, which collectively oscillate in anti-phase (Fig. 6). These two groups are separated by a thin layer of cells oscillating with the intermediate values of phase.

Clustering in the system with a large amount of elements exchanging chemical signals has become at the center of attention of many scientists recently. For example, a group of interacting with each other synthetic genetic oscillators has carefully been examined in a work [16]. Tissue clustering has been found to be divided into two types of oscillating cells in time. The authors of the work believe that this phenomenon is connected with two possible stable equilibria of the system. It is pointed out that the clustering is likely to be the most important characteristic of most communities and could be the reason of further cells differentiation in organs.

6 Conclusions

The research of the circadian rhythms synchronization at different scales of biological system description has been carried out on the basis of the suggested model of the circadian rhythms with time delay. The analysis of interaction at the scale of several genes has been done by stochastic modeling. Spatiotemporal dynamics of the protein concentration field at the scale of one or several cells has been studied within the deterministic model with diffusion. Finally, macroscopic effects arising at the community of a great amount of cells have been considered at the basis of the individual based chemo-mechanical model of the spreading epithelium, which cells could exchange chemical signals. The forms of the circadian rhythms synchronization at every scale of the system description have been identified.

We are grateful to L. M. Pismen and M. Salm agreed to hand over the software package to us to do the research in the area of complex living systems. The work was supported by the Department of Science and Education of Perm region (project C26/244), the Ministry of Science and Education of Russia (project 1.3103.2011) and Perm State Pedagogical University (project 031-F).

References

1. Pittendrigh, C.S.: Temporal organization: reflections of a Darwinian clock-watcher. Annu. Rev. Physiol. **55**, 16–54 (1993)
2. Stepanova, S.I.: Biorhythmological aspects of adaptation. Nauka, Moscow (1986)
3. Lakin-Thomas, P.L., Brody, S.: Circadian rhythms in microorganisms: new complexities. Annu. Rev. Microbiol. **58**, 489–519 (2004)
4. Pikovsky, A., Rosenblum, M., Kurths, J.: Synchronization—a universal concept in nonlinear sciences. Cambridge University Press, Cambridge (2001)
5. Hasty, J., Collins, J.J.: Translating the noise. Nat. Genet. **31**, 13–14 (2002)
6. Bratsun, D.A.: Effect of subcritical excitation of oscillations in stochastic systems with time delay. Part I. Regulation of gene expression. Comput. Res. Model. **3**, 421–438 (2011)
7. Ghaemmaghami, S., Won-Ki, H., Bower, K., Howson, R.W., Belle, A., Dephoure, N., O'Shea, E.K., Weissman, J.S.: Global analysis of protein expression in yeast. Nature **425**, 737–741 (2003)
8. Morozov, K.I., Pismen, L.M.: Cytoskeleton fluidization versus resolidification: prestress effect. Phys. Rev. E. **83**, 051920–051928 (2011)
9. Bratsun, D., Zakharov, A.: Modeling spatio-temporal dynamics of circadian rhythms in Neurospora crassa. Comput. Res. Model. **3**, 191–213 (2011)
10. Bratsun, D.A., Zakharov, A.P.: Deterministic modeling spatio-temporal dynamics of delay-induced circadian oscillations in Neurospora crassa. In: Interdisciplinary Symposium on Complex Systems, Czech Technical University, Prague, 10–13 Sept 2013
11. Smolen, P., Baxter, D.A., Byrne, J.H.: Modeling circadian oscillations with interlocking positive and negative feedback loops. J. Neurosci. **21**, 6644–6656 (2001)
12. Bratsun, D., Volfson, D., Hasty, J., Tsimring, L.S.: Delay-induced stochastic oscillations in gene regulation. Proc. Natl. Acad. Sci. U.S.A. **102**, 14593–14598 (2005)
13. Gillespie, D.T.: Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. **81**, 2340–2361 (1977)
14. Bratsun, D.A., Zakharov, A.P.: Adaptive numerical simulations of reaction-diffusion systems with history and time-delayed feedback. In: Interdisciplinary Symposium on Complex Systems, Czech Technical University, Prague, 10–13 Sept 2013
15. Salm, M., Pismen, L.M.: Chemical and mechanical signaling in epithelial spreading. Phys. Biol. **9**, 026009–026023 (2012)
16. Koseska, A., Ullner, E., Volkov, E., Kurths, J., Garcia-Ojalvo, J.: Cooperative differentiation through clustering in multicellular populations. J. Theor. Biol. **263**, 189–202 (2010)

Investigation on the Dynamics of PSO Algorithm Enhanced with Chaotic Lozi Map

Michal Pluhacek, Roman Senkerik, Ivan Zelinka and Donald Davendra

Abstract In this chapter, previously proposed utilization of discrete Lozi map based chaos pseudo-random number generator to enhance the performance of PSO algorithm is investigated with the detailed focus on the chaotic system dynamics. The elaborated tuning of chaotic system accessible parameters based experiment is presented here together with the investigation on the impact to the performance of PSO algorithm.

Keywords Particle swarm optimization · Swarm intelligence · Optimization · Chaos · Lozi map

1 Introduction

The particle swarm optimization algorithm (PSO) [1] is one of the best known and reputable swarm intelligence optimization algorithms which are a subset of evolutionary computation techniques (ECTs) [1–6]. These non-deterministic techniques are currently frequently combined with another nature inspired phenomenon – deterministic chaos [7]. Utilization of discrete chaotic systems as a chaotic pseudo-random

M. Pluhacek (✉) · R. Senkerik · I. Zelinka
Tomas Bata University in Zlin, Faculty of Applied Informatics, Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
e-mail: pluhacek@fai.utb.cz

R. Senkerik
e-mail: senkerik@fai.utb.cz

I. Zelinka
e-mail: zelinka@fai.utb.cz

D. Davendra
VŠB-Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, 17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
e-mail: donald.davendra@vsb.cz

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_35,

number generator (CPRNG) for the stochastic processes within the ECTs seems to have very positive effect on the performance of these techniques [8–18].

The performance of chaos driven ECTs such as Differential evolution [15, 16], PSO [9–14] and others [17, 18] was frequently investigated in recent studies. In this chapter, the PSO algorithm driven by chaotic Lozi map is investigated more in details with the focus on the influence of particular chaotic system setting and resulting chaotic dynamics.

Firstly, used PSO algorithm with inertia weight is explained. The next sections are aimed on the description of used chaotic system and experiments setup. Results and conclusion follow afterwards.

2 Particle Swarm Optimization Algorithm

The PSO algorithm is inspired by the natural swarm behavior of birds and fish. It was introduced by Eberhart and Kennedy in 1995 [1, 3] as an alternative to other ECTs, such as Ant Colony Optimization [2], Genetic Algorithms (GA) [4] or Differential Evolution (DE) [5]. Each particle in the population represents a possible solution of the optimization problem which is defined by its cost function. In each iteration, a new location (combination of cost function parameters) of the particle is calculated based on its previous location and velocity vector (velocity vector contains particle velocity for each dimension of the problem).

One of the disadvantages of the original PSO algorithm was poor local search ability. For this reason, several modifications of the PSO were introduced. The main principles of the PSO algorithm and its modifications are detailed in [1, 19, 20]. Within this research, the chaos driven PSO strategy with linear decreasing inertia weight was utilized [19, 20]. This strategy was first introduced in 1998 [19] in order to improve the local search capability of PSO. The selection of inertia weight strategy of PSO was based on numerous previous experiments [9, 10, 13, 14]. Default values of all PSO parameters were chosen according to the recommendations given in [1, 3, 19, 20].

Inertia weight is designed to influence the velocity of each particle differently over time [15, 16]. In the beginning of the optimization process, the influence of inertia weight factor w is minimal. As the optimization continues, the value of w is decreasing, thus the velocity of each particle is decreasing, since w is always the number less than one and it multiplies the previous velocity of particle in the process of new velocity value calculation. Inertia weight modification PSO strategy has two control parameters w_{start} and w_{end}. A new w for each iteration is given by (1), where i stand for current iteration number and n stands for the total number of iterations.

$$w = w_{start} - \frac{((w_{start} - w_{end}) \cdot i)}{n} \tag{1}$$

A chaos driven pseudo-random number generator is used in the main PSO formula (2) that determines new "velocity" and thus the position of each particle in the next iterations (or migration cycle).

$$v(i+1) = w \cdot v(i) + c_1 \cdot Rand \cdot (pBest - x(i)) + c_2 \cdot Rand \cdot (gBest - x(i)) \quad (2)$$

where:

$v(i+1)$ New velocity of a particle.
$v(i)$ Current velocity of a particle.
c_1, c_2 Priority factors.
pBest Best solution found by a particle.
gBest Best solution found in a population.
$x(i)$ Current position of a particle.
Rand Random number, interval (0, 1). Chaos number generator is applied only here.

The new position of a particle is then given by (3), where $x(i+1)$ represents the new position:

$$x(i+1) = x(i) + v(i+1) \quad (3)$$

3 Chaotic Lozi Map

The Lozi map is a simple discrete two-dimensional chaotic map. The Lozi map is depicted in Fig. 1. The map equations are given in (4). Typical parameters given in literature [7, 21, 22] are $a = 1.7$ and $b = 0.5$.

$$\begin{aligned} X_{n+1} &= 1 - a\,|X_n| + bY_n \\ Y_{n+1} &= X_n \end{aligned} \quad (4)$$

4 Experiment Setup

In Sect. 2 the chaos driven PSO algorithm with inertia weight was described. In Sect. 3 the Lozi map with typical control parameters was introduced. In this study the parameter a of the Lozi map was tuned and the influence of the changed chaotic system on the chaos driven PSO algorithm was analyzed. The first experiment was set as follows:

Population size $= 30$; *Iterations* $= 200$; *Dimension* $= 10$; *Lozi map parameter* $a = 1.5 - 1.7$ (step 0.1, 0.01, 0.001). For each setting the experiment was repeated 300 times and the mean result was recorded.

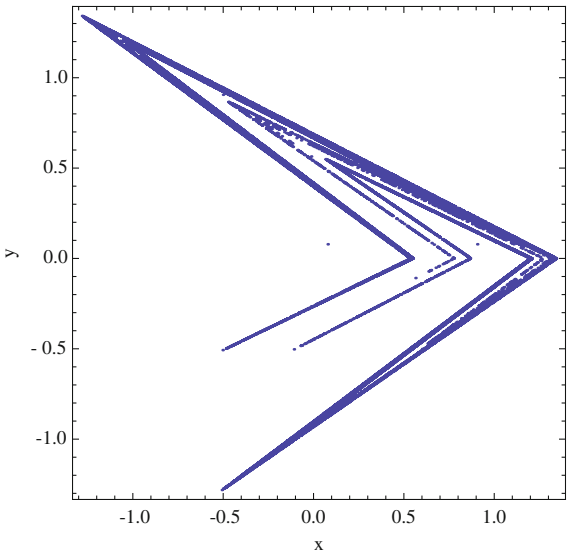


Fig. 1 x, y plot of Lozi map

4.1 Test Functions

Four different static test functions were used in this research: 1st De Jong's function (5), 2nd De Jong's function (6), Rastrigin's function (7) and Schwefel's function (8).

The 1st De Jong's function (Sphere function).

$$f(x) = \sum_{i=1}^{\dim} x_i^2 \tag{5}$$

Function minimum: Position for E_n: $(x_1, x_2 \ldots x_n) = (0, 0, \ldots, 0)$. Value for E_n: $y = 0$.

The 2nd De Jong's function (Rosenbrock's valley).

$$f(x) = \sum_{i=1}^{\dim} x_i^2 \tag{6}$$

Function minimum: Position for E_n: $(x_1, x_2 \ldots x_n) = (1, 1, \ldots, 1)$. Value for E_n: $y = 0$.

Rastrigin's function

$$f(x) = 10 \dim + \sum_{i=1}^{\dim} x_i^2 - 10 \cos(2\pi x_i) \tag{7}$$

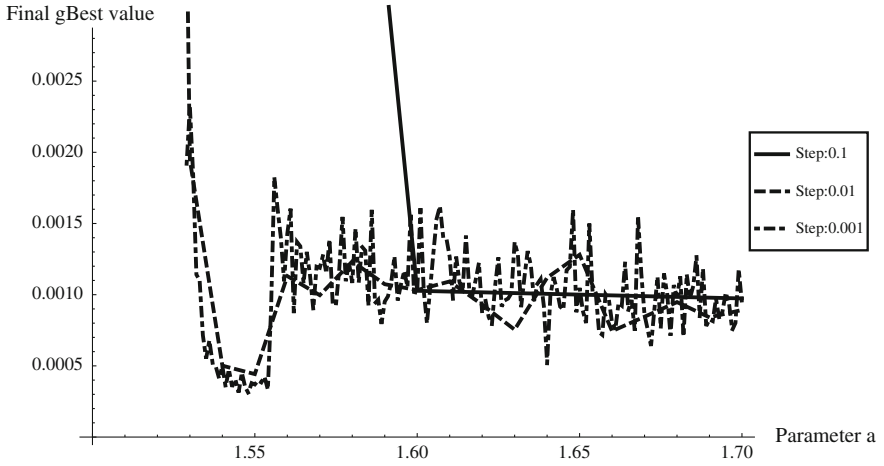


Fig. 2 Results of the tuning experiment—The 1st De Jong's function

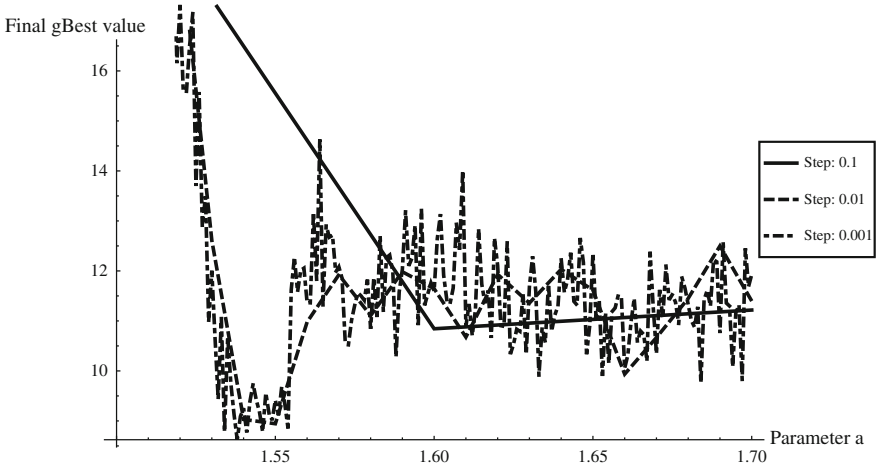


Fig. 3 Results of the tuning experiment—The 2nd De Jong's function

Function minimum: Position for E_n: $(x_1, x_2 \ldots x_n) = (0, 0, \ldots, 0)$. Value for E_n: $y = 0$.

Schwefel's function

$$f(x) = \sum_{i=1}^{\dim} -x_i \sin(\sqrt{|x|}) \tag{8}$$

Function minimum: Position for E_n : $(x_1, x_2 \ldots x_n) = (420.969, 420.969, \ldots, 420.969)$. Value for E_n: $y = -418.983 \cdot D$

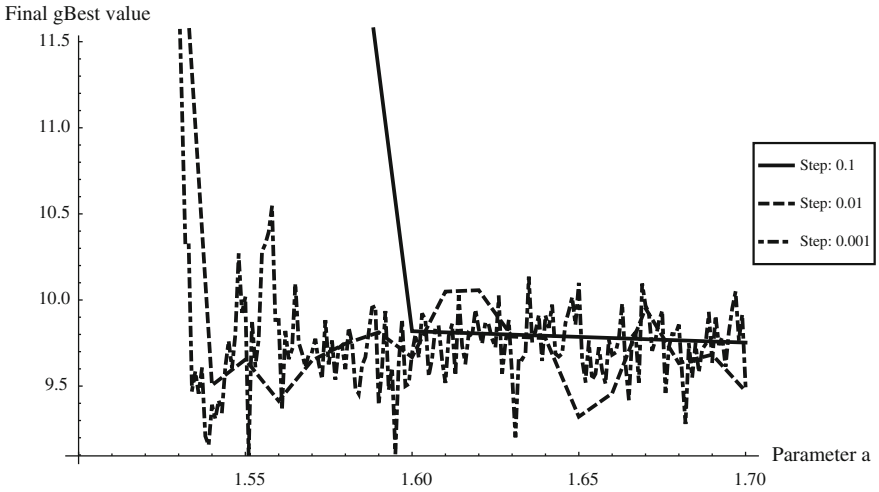


Fig. 4 Results of the tuning experiment—Rastrigin's function

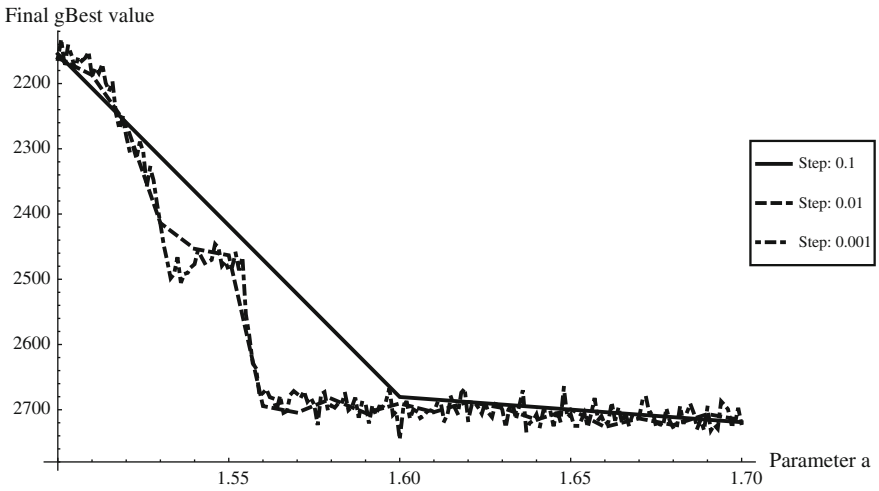


Fig. 5 Results of the tuning experiment—Schwefel's function

5 Results

The results obtained during the first experiment are depicted in Figs. 2, 3, 4 and 5. According to the obtained data the performance of chaos driven PSO algorithm on the 1st and 2nd De Jong's function seemed to be significantly better with Lozi map settings $a = 1.55$. (Figs. 2 and 3) in comparison with typically used value $a = 1.7$. Similar but less significant trend could be observed in Fig. 4 that depicts the results

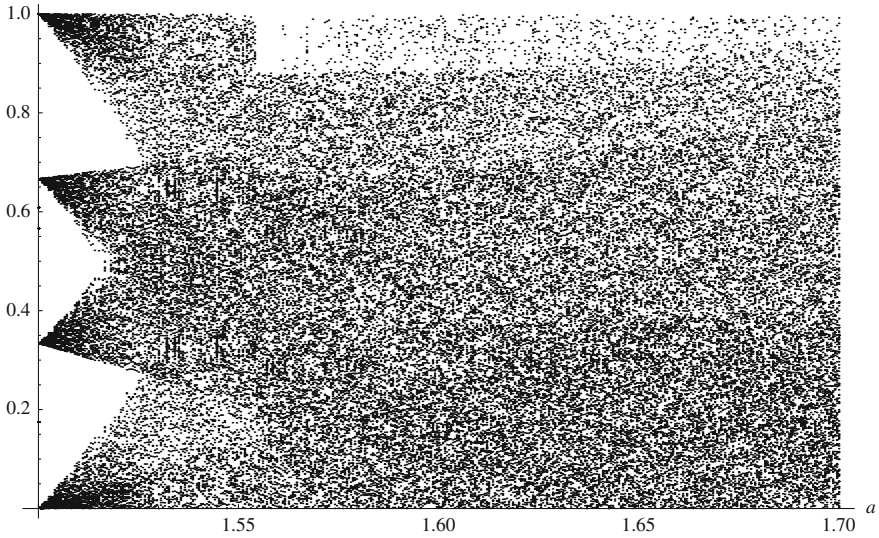


Fig. 6 Bifurcation diagram—Lozi map ($a = 1.5-1.7$, $b = 0.5$)

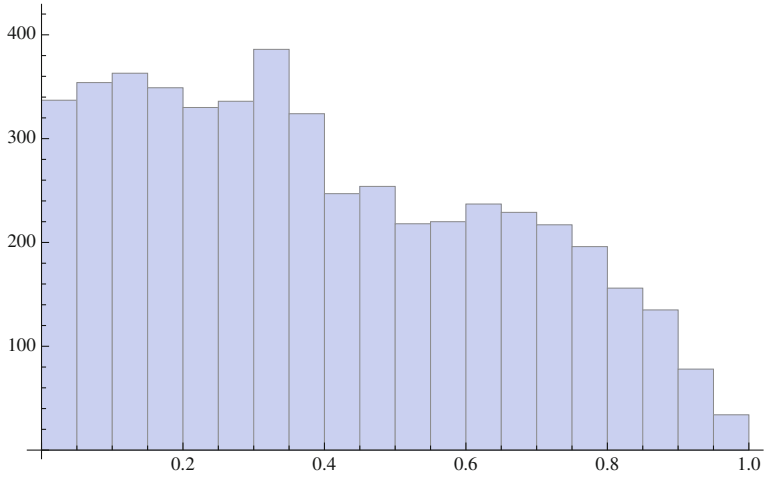


Fig. 7 Lozi map CPRNG distribution for $a = 1.7$

for Rastrigin's function. Surprisingly the totally opposite trend has arisen from the results of the first experiment with the Schwefel's function. For better understanding of this phenomenon, Fig. 6 depicts the bifurcation diagram for chaotic Lozi map for fixed $b = 0.5$ and $a = 1.5$–1.7. Furthermore the distributions of CPRNGs based on Lozi map with different a settings are presented on Figs. 7 and 8.

After the first experiment focused on the detailed investigation on the influence of Lozi map accessible parameter tuning subsequently the performance testings and

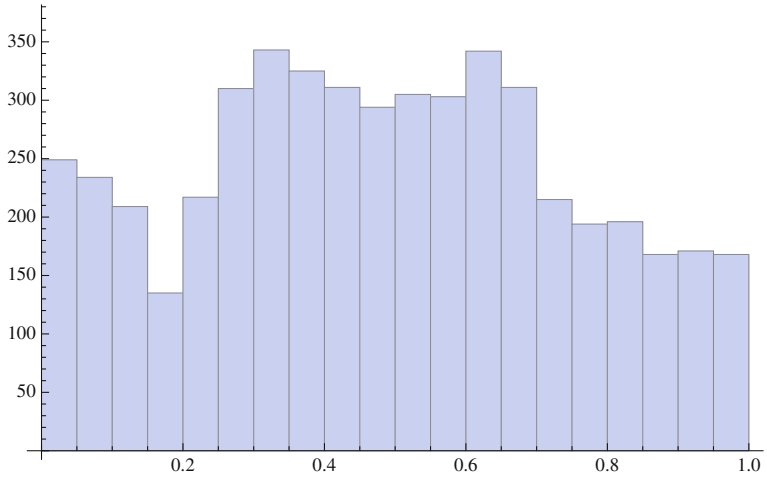


Fig. 8 Lozi map CPRNG distribution for $a = 1.55$

Table 1 Mean results of the second experiment—The 1st De Jong's function

DIM	20	40	60	80	100
PSO Weight	0.31126	3.26528	7.74353	12.6881	18.3695
PSO Lozi	0.48832	4.66315	10.7448	18.3021	26.2179
PSO Lozi 2	0.15302	1.97864	4.75604	8.32826	11.8904

Table 2 Mean results of the second experiment—The 2nd De Jong's function

DIM	20	40	60	80	100
PSO Weight	89.1357	585.711	1488.26	2686.2	4093.41
PSO Lozi	112.415	761.655	1993.22	3595.95	5483.92
PSO Lozi 2	60.0585	330.803	774.254	1380.46	2051.72

Table 3 Mean results of the second experiment—Rastrigin's function

DIM	20	40	60	80	100
PSO Weight	34.6047	133.86	258.134	393.647	543.166
PSO Lozi	38.1713	145.831	272.309	418.132	569.274
PSO Lozi 2	32.4471	124.82	239.015	367.258	510.703

comparisons for the chaos driven PSO algorithm were carried out. Following Tables 1, 2, 3 and 4 contain mean results of this experiment. The notation is: PSO Weight—PSO with non-chaotic PRNG, PSO LOZI – PSO with Lozi map based CPRNG (a = 1.7) and PSO Lozi 2 – PSO with Lozi map based CPRNG (a = 1.55). The experiment was repeated 100 times for different benchmark functions dimension setting. The

Table 4 Mean results of the second experiment—Schwefel's function

DIM	20	40	60	80	100
PSO Weight	−3846.19	−6075.94	−8060.01	−9829.3	−11339.1
PSO Lozi	−4282.11	−6657	−8450.79	−10075.8	−11321.5
PSO Lozi 2	−3867.85	−6296.24	−8140.84	−9748.96	−11159.8

PSO algorithm was set as follows: Population size: 30; Iterations: $20 \cdot dim$. The bold values within the Tables 1, 2, 3 and 4 depict the best obtained solutions.

6 Conclusion

In this chapter, the enhanced PSO algorithm performance with Lozi map CPRNG was investigated in details. The accessible parameters of chaotic system were tuned and resulting impact of different setting of the system on the performance of chaos driven PSO algorithm was recorded and analyzed. Four different test functions were used for all experiments within this research. These test functions were utilized for the performance comparisons of the PSO algorithm with CPRNG with setting based on the fine tuning of Lozi map against the both canonical PSO with inertia weight and PSO algorithm that utilized CPRNG typical setting from literature and previous successful experiments with the concept of embedding the chaotic dynamics into the evolutionary techniques.

Obtained results lend weight to the importance of previous research focused on the embedding of CPRNG into the evolutionary techniques and at the same time these results represents the next step in this research, which is the fine tuning of chaotic maps parameters used as CPRNG.

The presented data seem to indicate that the proper setting of chaotic system that is used as a CPRNG for PSO algorithm could lead to significant improvement of the performance of such algorithm even for different test functions and dimension value settings.

Since this is an initial study, the deeper statistical analyses based on paired t-tests or parametric tests will be performed soon within the future research. The *C* language "*Mersenne Twister*" with default automatic setting pseudo-random number generator was applied to represent traditional pseudo-random number generators in comparisons.

Acknowledgments This work was supported by Grant Agency of the Czech Republic-GACR P103/13/08195S, by the project Development of human resources in research and development of latest soft computing methods and their application in practice, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational Program Education for Competitiveness, co-financed by ESF and state budget of the Czech Republic; European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089, by the Technology Agency of the Czech Republic under the

Project TE01020197, and by Internal Grant Agency of Tomas Bata University under the projects No. IGA/FAI/2013/012.

References

1. Kennedy, J., Eberhart, R., Particle swarm optimization. Proceedings of IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
2. Dorigo, M.: Ant Colony Optimization and Swarm Intelligence. Springer, Berlin (2006)
3. Kennedy, J., Eberhart, R.C., Shi, Y.: Swarm Intelligence. The Morgan Kaufmann Series in Artificial Intelligence. Morgan Kaufmann, San Francisco (2001)
4. Goldberg, D.E., Genetic Algorithms in Search Optimization and Machine Learning. p. 41. Addison Wesley (1989). ISBN 0201157675
5. Storn, R., Price, K.: Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. **11**, 341–359 (1997)
6. Zelinka, I., SOMA—self organizing migrating algorithm. In: Babu, B.V., Onwubolu, G. (eds.) New Optimization Techniques in Engineering, vol. 33, Springer-Verlag (2004). ISBN: 3-540-20167X
7. Caponetto, R., Fortuna, L., Fazzino, S., Xibilia, M.G.: Chaotic sequences to improve the performance of evolutionary algorithms. IEEE Trans. Evol. Comput. **7**(3), 289–304 (2003)
8. Davendra, D., Zelinka, I., Senkerik, R.: Chaos driven evolutionary algorithms for the task of PID control. Comput. Math. Appl. **60**(4), 1088–1104 (2010). ISSN 0898–1221
9. Pluhacek, M., Senkerik, R., Davendra, D., Zelinka, I.: Designing PID controller for DC motor system by means of enhanced PSO algorithm with discrete chaotic Lozi map. In: Proceedings of the 26th European Conference on Modelling and Simulation, ECMS 2012, pp. 405–409 (2012). ISBN 978-0-9564944-4-3
10. Pluhacek, M., Senkerik, R., Davendra, D., Zelinka, I.: PID controller design for 4th order system by means of enhanced PSO algorithm with Lozi chaotic map. In: Proceedings of the 18th International Conference on Soft Computing MENDEL 2012, pp. 35–39. (2012). ISBN 978-80-214-4540-6
11. Araujo, E., Coelho, L.: Particle swarm approaches using Lozi map chaotic sequences to fuzzy modelling of an experimental thermal-vacuum system. Appl. Soft Comput. **8**(4), 1354–1364 (2008)
12. Alatas, B., Akin, E., Ozer, B.A.: Chaos embedded particle swarm optimization algorithms. Chaos, Solitons Fractals **40**(4), 1715–1734 (2009). ISSN 0960–0779
13. Pluhacek, M., Senkerik, R., Davendra, D., Kominkova Oplatkova, Z., Zelinka I.: On the behavior and performance of chaos driven PSO algorithm with inertia weight. Comput. Math. Appl. **66**, 122–134 (2013). ISSN 0898–1221
14. Pluhacek, M., Budikova, V., Senkerik, R., Oplatkova, Z., Zelinka, I.: On the performance of enhanced PSO algorithm with Lozi chaotic map—an initial study. In: Proceedings of the 18th International Conference on Soft Computing MENDEL 2012, pp. 40–45. (2012). ISBN 978-80-214-4540-6
15. Senkerik, R., Davendra, D., Zelinka, I., Pluhacek, M., Oplatkova, Z.: An investigation on the differential evolution driven by selected discrete chaotic systems. In: Proceedings of the 18th International Conference on Soft Computing MENDEL 2012, pp. 157–162. (2012). ISBN 978-80-214-4540-6
16. Ozer, A.B.: CIDE: Chaotically initialized differential evolution. Expert Syst. Appl. **37** (6), 4632–4641 (2010). ISSN 0957–4174
17. Alatas, B.: Chaotic bee colony algorithms for global numerical optimization. Expert Syst. Appl. **37**(8), 5682–5687 (2010). ISSN 0957–4174
18. Gandomi, A.H., Yang, X.S., Talatahari, S., Alavi, A.H.: Firefly algorithm with chaos. Commun. Nonlinear Sci. Numer. Simul. **18**(1), 89–98 (2013). ISSN 1007–5704

19. Shi, Y.H., Eberhart, R.C.: A modified particle swarm optimizer. In: Proceedings of IEEE International Conference on Evolutionary Computation, pp. 69–73 (1998)
20. Nickabadi, A., Ebadzadeh, M.M., Safabakhsh, R.: A novel particle swarm optimization algorithm with adaptive inertia weight. Appl. Soft Comput. **11**(4), 3658–3670 (2011). ISSN 1568–4946
21. Sprott, J.C.: Chaos and Time-Series Analysis. Oxford University Press, Oxford (2003)
22. Aziz-Alaoui, M.A., Robert, C., Grebogi, C.: Dynamics of a Hénon-Lozi-type map. Chaos, Solitons Fractals **12**, 2323–2341 (2001). ISSN 0960–0779

On the Development of Complex Cost Function for the Evolutionary Chaos Control: A Brief Study

Roman Senkerik, Ivan Zelinka, Michal Pluhacek, Zuzana Kominkova Oplatkova and Roman Jasek

Abstract This work represents the brief introduction into the issue of development of complex cost function for evolutionary optimization of control of discrete chaotic systems. This work introduces briefly the evolutionary approach representing tuning of parameters for an existing control method. The main part of this work is focused on the process of development of the proper cost function design used within the evolutionary process. As an example of discrete chaotic system, two-dimensional Hénon map was used.

Keywords Chaos control · Optimization · Evolutionary algorithms

1 Introduction

During the recent years, usage of new soft-computing techniques in engineering, technology, modeling, computing and simulations has attracted the attention of researchers worldwide. Presently, evolutionary algorithms are known as a powerful set of tools for almost any difficult and complex optimization problem.

The interest about the interconnection between evolutionary techniques and control of chaotic systems is spread daily. The first steps were done in [1–3], where the control law was based on the Pyragas method, which is Extended delay feedback control (ETDAS) [4–6] for the purpose of stabilization of a chaotic systems on desired Unstable Periodic Orbits (UPO). This method is very advantageous for

R. Senkerik · I. Zelinka · M. Pluhacek · Z. Kominkova Oplatkova (✉) · R. Jasek
Tomas Bata University in Zlin, Faculty of Applied Informatics, Nam T.G. Masaryka 5555, 760 01 Zlin, Czech Republic
e-mail: kominkovaoplatkova@fai.utb.cz

I. Zelinka
Technical University of Ostrava, Faculty of Electrical Engineering and Computer Science, 17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_36,

evolutionary computation, due to the amount of accessible control parameters, which can be easily tuned by means of evolutionary algorithms (EA).

The main core of the optimization problem is the proper definition of the cost function (also called fitness function). It represents the quality of solution based on the combination of input parameters and given constraints within the n-dimensional space.

This work is a cumulation of previous research [7–9] and expansion of experiences from the initial studies [10, 11]. Furthermore, this work briefly introduces the background of the research, mainly the development of the cost function and show general characteristics of obtained results.

2 Problem Design

The brief description of used chaotic system, original feedback chaos control method ETDAS is given here. The process of development of complex cost functions follows in the next chapter.

2.1 Selected Chaotic System

The chosen example of a chaotic system used for generating of illustrative cost function surfaces and simulation outputs was the two dimensional Hénon map in form (1):

$$\begin{aligned} x_{n+1} &= a - x_n^2 + by_n \\ y_{n+1} &= x_n \end{aligned} \tag{1}$$

The Hénon map is a discrete-time dynamical system. The map depends on two parameters, a and b, which for the canonical Hénon map have values of $a = 1.4$ and $b = 0.3$. For these canonical values the Hénon map is chaotic [12].

2.2 ETDAS Control Method

This work is focused on the explanation of cost function design for tuning of parameters for ETDAS method control law to stabilize desired Unstable Periodic Orbits (UPO). In the described research, desired UPOs were p-1 (stable state) and p-2 (two-periodic orbit, which represents oscillation between two values). The original control method—ETDAS in the discrete form suitable for Hénon map has the form (2).

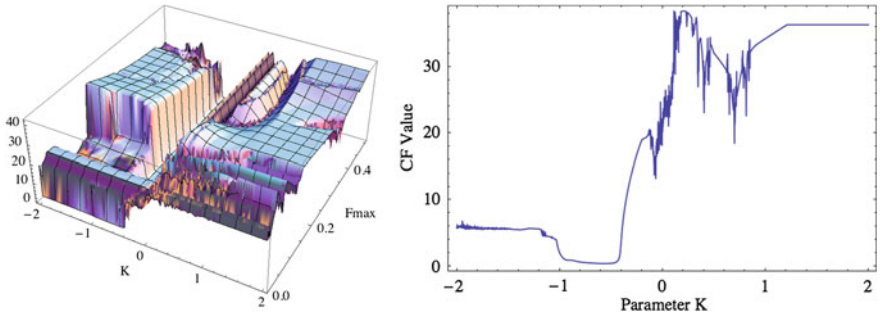


Fig. 1 Dependence of CF value on parameters K and $F_{\max}$ (*left*), and on parameter K (*right*); Simple cost function, p-1 orbit

$$\begin{aligned} x_{n+1} &= r x_n \left(1 - x_n\right) + F_n \\ F_n &= K\left[\left(1 - R\right) S_{n-m} - x_n\right] \\ S_n &= x_n + R S_{n-m} \end{aligned} \tag{2}$$

where: K and R are adjustable constants, which has to be evolutionary tuned. F is the perturbation; S is given by a delay equation utilizing previous states of the system, m is the period of m-periodic orbit to be stabilized. The perturbation F_n in Eqs. (2) may have arbitrarily large value, which can cause diverging of the system outside the output interval of Logistic equation $\{0, 1\}$. Therefore, F_n should have a value between $< -F_{\max}, F_{\max} >$. The suitable $F_{\max}$ value was also obtained from evolutionary optimization process.

3 Cost Functions Development

The proposal of the basic cost function (CF) is in general based on the simplest CF, which could be used problem-free only for the stabilization of p-1 orbit. The idea was to minimize the area created by the difference between the required state and the real system output on the whole simulation interval—τ_i. This CF design is very convenient for the evolutionary searching process due to the relatively favorable CF surface. (See Fig. 1.) Nevertheless this simple approach has one big disadvantage, which is the including of initial chaotic transient behavior of not stabilized system into the cost function value. As a result of this, the very tiny change of control method setting for extremely sensitive chaotic system causing very small change of CF value, can be suppressed by the above-mentioned including of initial chaotic transient behavior

Another cost function had to be used for stabilizing of the higher periodic orbit, due to simple and serious reason, which is the degrading of the possible best solution

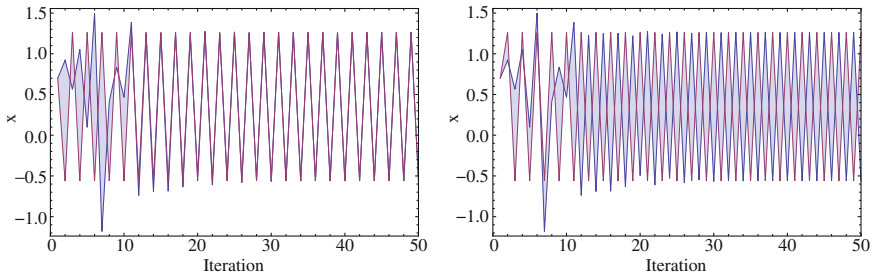


Fig. 2 The example of possible phase shift of periodic orbit during evolutionary process: No UPO shift—good solution (*left*), Full phase shift—Even though perfectly stabilized, for EA it is "bad" solution (*right*); Light blue area represents the difference between output of the system (*blue line*) and the required behavior (*red line*), p-2 orbit

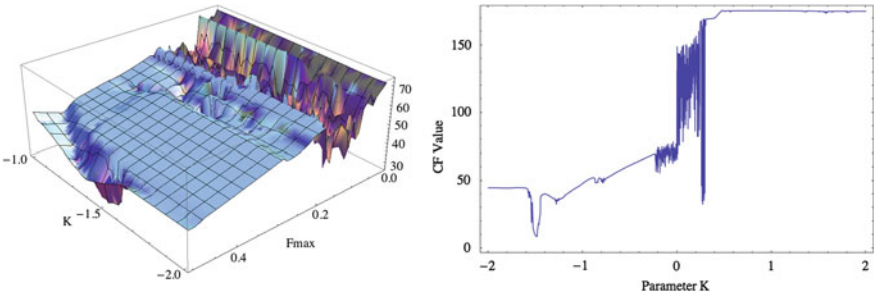


Fig. 3 Dependence of CF value on parameters K and F_{max} (*left*), and on parameter K (*right*); Simple cost function with selection rule, p-2 orbit. The global optimum ($K = -1.48$) represents solution with very fast but not precise (or only temporary) stabilization, whereas the solution given by local optimum ($K = 0.34$) secures slow but very precise stabilization

by a phase shift of periodic orbit. The example of possible phase shift is depicted in Fig. 2.

Such a cost function is based on the minimization of the area created by the difference between the all possible required states and the real system output on the whole simulation interval. In case of the p-2 orbit, during the CF evaluation, the difference between system output and the two possible required behaviors (i.e. oscillations A-B-A-B… or B-A-B-A) is computed and thereafter selected the minimal one as a CF output (3). An interesting phenomenon has occurred in this case. The global optimum represents solution with very fast but not precise (or only temporary) stabilization, whereas the solution given by local optimum secures slow but very precise stabilization on desired higher periodic orbit. This problem is caused by including of initial chaotic part into CF value before stabilization together with the increasing of the nonlinear and chaotic character of CF surface due to the selection rule within the CF evaluation. The 2D plot of dependence of CF values on one selected optimized parameter and 3D CF surface is depicted in Fig. 3.

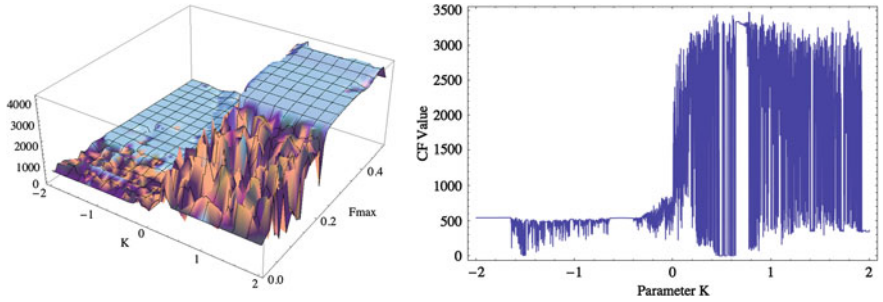


Fig. 4 Dependence of CF value on parameters K and F_{max} (*left*), and on parameter K (*right*); Simple cost function with selection rule and penalization, p-2 orbit

$$CF_{SELECT} = Min \left\{ \begin{array}{l} \sum\limits_{t=0}^{\tau i} |TS1_t - AS_t| \\ \\ \sum\limits_{t=0}^{\tau i} |TS2_t - AS_t| \end{array} \right\} \tag{3}$$

where:

AS actual state
TS1 target state 1 (i.e. oscillations A-B-A-B...)
TS2 target state 2 (i.e. oscillations B-A-B-A...)
τ_i simulation interval

The issue of unsuitable CF surface with global optimum representing solution with very fast but not precise (or only temporary) stabilization, and with local optimum solutions securing slow but very precise stabilization, can be easily suppressed by simple penalization within CF evaluation. The difference between system output and the two possible required behaviors (i.e. oscillations A-B-A-B… or B-A-B-A) evaluated on the whole simulation interval, as suggested in (3), is multiplied by the similar difference, but evaluated only from the last 20 iterations. In the case of the temporary or not precise stabilization, such a difference will be represented by a number mostly higher than 1. This means, that the solution will be penalized. In the opposite case of very precise stabilization, the difference will be very close to zero, so that the final solution will be more favorable in case of the minimizing evolutionary optimization. Such a CF design is given in (4) and the 3D CF surface is depicted in Fig. 4.

$$CF_{SELECT\_P} = Min \left\{ \begin{array}{l} \sum\limits_{t=0}^{\tau i} |TS1_t - AS_t| \cdot \sum\limits_{t=\tau_i-20}^{\tau i} |TS1_t - AS_t| \\ \\ \sum\limits_{t=0}^{\tau i} |TS2_t - AS_t| \cdot \sum\limits_{t=\tau_i-20}^{\tau i} |TS2_t - AS_t| \end{array} \right\} \tag{4}$$

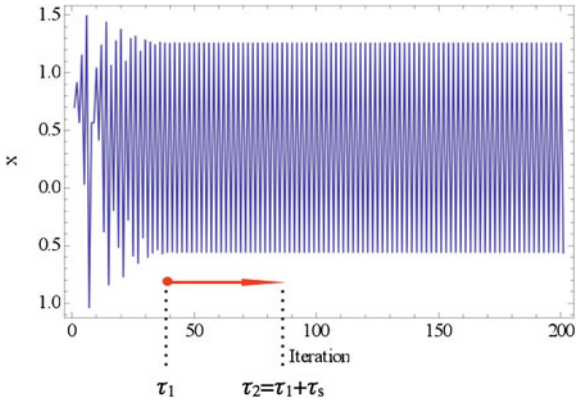


Fig. 5 "Floating window" for minimization

where:

AS actual state
TS1 target state 1 (i.e. oscillations A-B-A-B. . .)
TS2 target state 2 (i.e. oscillations B-A-B-A. . .)
τ_i simulation interval

The penalization, which is given by multiplication of two differences between target state and actual output state of chaotic system, causes increase of nonlinearity and chaotic nature of CF surface. Although the issue of problems with global/local solutions was suppressed by this CF design (4), another issue makes this CF design less convenient (or comfortable to use) for numerous experiments with different chaotic systems and different desired UPOs. This CF design requires preparing (programming) of unique functions for different UPOs and systems. Thus the different universal CF (5) served as a core for advanced CF used within advanced evolutionary approach for chaos control optimization.

Different type of universal cost functions without any selection rules are purely based on searching for the desired stabilized periodic orbit and thereafter calculation of the difference between desired and found actual periodic orbit on the short time interval—τ_s (20 iterations—p-1 orbit and 40 iterations—p-2 orbit) from the point, where the first minimal value of difference between desired and actual system output is found (i.e. floating window for minimization—see Fig. 5.).

Such a design of universal CF should secure the successful stabilization of either p-1 orbit (stable state) or any else higher periodic orbit anywise phase shifted.Furthermore, due to CF values converging towards zero, this CF also allows the using of decision rules, avoiding very time demanding simulations. This rule stops EA immediately, when the first individual with good parameter structure is reached, thus the value of CF is lower then the acceptable (CF_{acc}) one. Based on the numerous experiments, typically $CF_{acc} = 0.001$ at time interval $\tau_s = 20$ iterations, thusthe difference between desired and actual output has the value of 0.0005 per

iteration—i.e. successful stabilization for the used control technique. The CF_{UNI} has the form (5).

$$CF_{UNI} = pen_1 + \sum_{t=\tau 1}^{\tau 2} |TS_t - AS_t| \tag{5}$$

where:

TS	target state, AS—actual state
τ_1	the first min value of difference between TS and AS
τ_2	the end of optimization interval ($\tau_1 + \tau_s$),
$pen_1 = 0$	if $\tau_i - \tau_2 \geq \tau_s$, $pen_1 = 10^*(\tau_i - \tau_2)$ if $\tau_i - \tau_2 < \tau_s$ (i.e. late stabilization).

For the development of advanced CF securing the very fast stabilization, it was necessary to modify the definition of universal CF_{UNI} in order to decrease the average number of iteration required for the successful stabilization and avoidance of any associated problem. The easiest but the most problematic way is that the whole CF value is multiplied by the number of iterations (*NI*) of the first found minimal value of difference between desired and actual system output (i.e. the beginning of fully stabilized UPO). To avoid errors associated with CF returning value 0 and other problems, the small constant (*SC*) is added to CF value before penalization (multiplying by *NI*). The *SC* value (7) is computed with the aid of power of non-penalized basic part of CF (6), thus it is always secured that the penalization is at similar level as the non-penalized CF value.

$$\text{ExpCF} = \log_{10}\left(\sum_{t=\tau 1}^{\tau 2} |TS_t - AS_t| + 10^{-15}\right) \tag{6}$$

$$SC = 10^{\text{ExpCF}} \tag{7}$$

In general, there exists two possible ways for applying the multiplication by the number of iterations required for stabilization (*NI*). The first version of the final design of targeting CF (CF_{TARG1}) has the form (8). Here the sum of basic part of CF and automatically computed *SC* is multiplied by *NI*. Consequently, the EA should find the solutions securing the fast targeting into desired behavior of the system.

$$CF_{TARG1} = NI\left(SC + p1 + \sum_{t=\tau 1}^{\tau 2} |TS_t - AS_t|\right) \tag{8}$$

In the second version of targeting CF (CF_{TARG2}), there is only a slight change in the comparison with the previous proposal. The number of steps for the stabilization (*NI*) multiplies only with the small constant (*SC*), which is counted in the same way as in the previous case (12). This version of targeting CF (CF_{TARG2}) has the form (9).

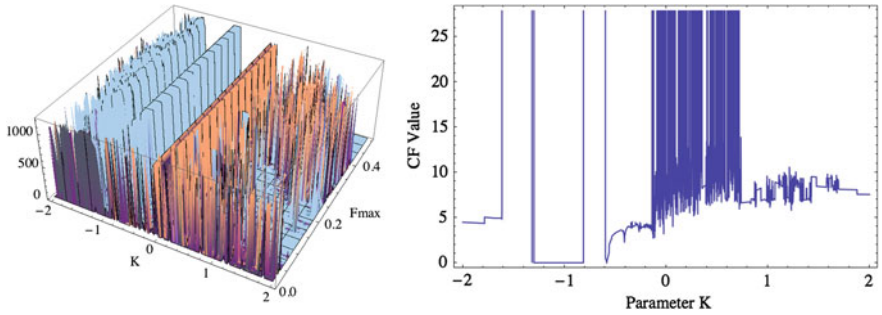


Fig. 6 Dependence of CF value on parameters K and $F_{\max}$ (*left*), and on parameter K (*right*); Advanced cost function, for simple evolutionary process, p-1 orbit

Table 1 Simulation results characteristics

Cost function	Advantages	Disadvantages
CF_{SIMPLE}	Wide global extreme area, simple and fast evolutionary process	Suitable only for p-1 orbit optimization case, no constraints or penalizations could be added
CF_{SELECT}	Suitable for any UPO, quite favorable CF surface for evolutionary process	Problems with global/local extremes, for different UPO order, the CF has to be re-designed, no constraints or penalizations could be added
CF_{SELECT_P}	Suitable for any UPO, guarantee of precise UPO stabilization, best solution is very close to the CF value zero.	Increased nonlinearity and chaotic nature of CF surface, for different UPO order, the CF has to be re-designed
CF_{UNI}	Universal CF suitable for any UPO, arbitrary constraints or penalizations could be simply added	High nonlinearity and chaotic nature of CF
CF_{TARG1}/CF_{TARG2}	Universal CF suitable for any UPO with inbuilt optimization of time required for full stabilization	Extremely high nonlinearity and chaotic (discrete) nature of CF

$$CF_{TARG2} = (NI \cdot SC) + p1 + \sum_{t=\tau 1}^{\tau 2} |TS_t - AS_t| \tag{9}$$

The issue of pure searching for periodic orbits causes very chaotic, erratic and discrete type CF surfaces (see Fig. 6.)

4 Results Overview

This section contains the brief overview of the simulation results characteristics given as the advantages/disadvantages for each single cost function design depicted in Table 1.

5 Conclusion

This work introduces the possible approach for the optimization of stabilization of Hénon map, which was selected as an example of discrete chaotic system. The main part of this work is focused on the issue of developing the proper cost function for evolutionary searching process.

The question of energy costs and more precise stabilization will be included into future research together with the development of better cost functions, and performing of numerous simulations to obtain more results and produce better statistics, thus to confirm the robustness of this approach.

Acknowledgments This work was supported by Grant Agency of the Czech Republic- GACR P103/13/08195S, by the project Development of human resources in research and development of latest soft computing methods and their application in practice, reg. no. CZ.1.07/2.3.00/20.0072 funded by Operational Program Education for Competitiveness; by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089, and by Internal Grant Agency of Tomas Bata University under the project No. IGA/FAI/2013/012.

References

1. Zelinka, I., Senkerik, R., Navratil, E.: Investigation on evolutionary optimization of chaos control. Chaos, Solitons Fractals **40**(1), 111–129 (2009)
2. Senkerik, R., Zelinka, I., Davendra, D., Oplatkova, Z.: Evolutionary design of chaos control in 1D. In: Zelinka, I., Celikovski, S., Richter, H., Chen, G. (eds.) Evolutionary Algorithms and Chaotic Systems, pp. 165–190. Springer-Verlag, Berlin (2010)
3. Senkerik, R., Zelinka, I., Davendra, D., Oplatkova, Z.: Utilization of SOMA and differential evolution for robust stabilization of chaotic Logistic equation. Comput. Math. Appl. **60**(4), 1026–1037 (2010)
4. Pyragas, K.: Control of chaos via extended delay feedback. Phys. Lett. A **206**, 323–330 (1995)
5. Just, W.: Principles of time delayed feedback control. In: Schuster, H.G. (eds.) Handbook of Chaos Control, Wiley-Vch, Weinheim (1999)
6. Pyragas, K.: Continuous control of chaos by self-controlling feedback. Phys. Lett. A **170**, 421–428 (1992)
7. Senkerik, R., Zelinka, I., Davendra, D., Oplatkova, Z.: Evolutionary optimization of Henon map control–a Blackbox approach. Int. J. Oper. Res. **13**(2), 129–146 (2012)
8. Senkerik, R., Oplatkova, Z., Zelinka, I., Jasek, R.: Application of Analytic Programming for Evolutionary Synthesis of Control Law—Introduction of Two Approaches. In: Byrski, A., Oplatkova, Z., Carvalho, M., Kisiel-Dorohinicki, M. (eds) Advances in Intelligent Modelling

and Simulation: Studies in Computational Intelligence. Springer Berlin Heidelberg, vol 416, pp 253–268 (2012) doi:10.1007/978-3-642-28888-3_10
9. Senkerik, R., Oplatkova, Z., Zelinka, I., Davendra, D., Jasek, R.: Application of evolutionary techniques for optimization of Chaos control—introduction of three approaches. In: Zelinka, I., Snasel, V., Abraham, A. (eds.) Handbook of Optimization, vol 38. Intelligent Systems Reference Library. Springer Berlin Heidelberg, pp 801–820 (2013) doi:10.1007/978-3-642-30504-7_31
10. Senkerik R, Oplatkova Z, Zelinka I, Davendra D (2013) Synthesis of feedback controller for three selected chaotic systems by means of evolutionary techniques: Analytic programming. Mar. Technol. Soc. 57(1–2), 57–67. doi:10.1016/j.mcm.2011.05.030
11. Kominkova, Oplatkova Z., Senkerik, R., Zelinka, I., Pluhacek, M.: Analytic programming in the task of evolutionary synthesis of a controller for high order oscillations stabilization of discrete chaotic systems. Comput. Math. Appl. **66**(2013), 177–189 (2013)
12. Hilborn, R.C.: Chaos and Nonlinear Dynamics: An Introduction for Scientists and Engineers. Oxford University Press, Oxford (2000)

4-D Seismic Tomography for the Complex System of Strong Earthquakes: Formulation of a Problem

Tatyana A. Smaglichenko and Ingi Th. Bjarnason

Abstract Geodynamic processes are acting in the Earth's interior and they cause earthquakes of various intensity. Earthquakes occur randomly and they are often in clusters. Sometimes it happens that before strong earthquakes there is a seismic quiescence that is characterized by the absence of significant seismic events. This may indicate that Earth's geological system prepares itself for a catastrophe. Complexity theory describes regularities of the behavior of dynamical systems before the occurrence of a disaster. The main part of this chapter is formulating a problem to investigate the behavior of a geophysical parameter, namely seismic velocity before the occurrence of the strong earthquake. Considering that velocity is a random variable, we apply the distribution function to estimate the dynamic state of the strong earthquakes complex system.

Keywords Seismic tomography · Velocity model · Statistics · Geodynamics

1 Introduction

The Earth is a dissipative system of heterogeneous geological structures. The state of structures can be changed under the influence of various physical and chemical factors. This leads e.g. to the unpredictable occurrences of strong earthquakes. In his work Nicolis and Prigogine [8] examined various states of stability of dissipative systems. The states can be described via the Lyapunov function [6]. However this is not always possible. For example, it is difficult to describe the un-insulated

T. A. Smaglichenko (✉)
Research Oil and Gas Institute of Russian Academy of Sciences, Moscow, Russia
e-mail: t.a.smaglichenko@gmail.com

I. Th. Bjarnason
Institute of Earth Sciences, University of Iceland, Reykjavík, Iceland
e-mail: ingib@hi.is

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_37,

systems via an ideal function. According to Nicolis and Prigogine [8] there are several important control parameters. Critical values of these parameters are responsible for the instability of the system. Once parameters pass their critical values the system enhances fluctuations and the system can cause a catastrophic state.

Nowadays various seismic parameters and measurements are used in order to investigate a behavior of the complex system of strong earthquakes with the aim to predict strong earthquakes. Forecast models are constructed using catalog information, which contain magnitudes of earthquakes prior and after the main shock. In this approach the time window, within which a forecast is made, is often selected by taking into account information on historical and pre-historical earthquakes that occurred in the region of study [7, 16]. Other factors to determine of the forecast time window are based on the seismic activity that is instrumentally recorded in a region. Then forecasts are given for 1 day up to couple of years depending on the detailed information in the seismic catalog data [4, 15]. The probability and magnitude of a catastrophic event can be estimated by studying complex earthquake networks and constructing the so called chain graphs [1].

The other line of approach in studying the earthquake forecast problem is based on determining the physical parameters and the wider geodynamical effects of the earthquake source. The response of an elastic medium to the seismic wave propagation is analyzed. One of the main observations for this analysis is the polarity of the first arriving P-wave as a function of azimuth. From the P-wave polarity distribution it is possible to estimate the earthquake mechanism. The stress/strain fields in the Earth's crust are estimated by inverting the earthquake mechanisms. For example it was shown that after the 2011 Tohoku earthquake ($M = 9.1$) in Japan, the main axis of the stress field was rotated 30–35° with respect to Pacific plate boundary [3]. In order to map in detail the stress field, Rebetskii [9, 10] has developed the method of cataclastic analysis of discontinuous displacement. It has been shown that the fault rupture in the 2004 Sumatra–Andaman earthquake ($M = 9.1$), the 2003 Tokachi–Ochi earthquake ($M = 8.3$), the 2006 Simushir earthquake ($M = 8.3$), and the 2010 Chile Earthquake ($M = 8.3$) was started on the border of areas of high and low pressure levels, i.e. in areas with high pressure gradient [10]. These results have an important implication in understanding the likely geographical location of large earthquakes.

The second main parameter to consider is a time. It is already established that accumulation of stresses in the crust is characterized by time cycles. Cycles can be estimated by measuring movements of the crust with GPS data (both the sudden movement in earthquakes and the gradual drift of the crust) and by account of historical seismicity and paleoseismical data that can be determined with methods of field-geology. In a region a "normal" cycle may e.g. be defined for earthquakes with magnitudes 7 and less and a "mega" cycle for earthquakes with magnitudes greater than 8. In other regions "mega" cycle may be defined to consist of earthquakes of magnitude ~9. Careful evaluation of the geological field data suggests that the 2011 Tohoku earthquake may have been part of a 500-year-long cycle of "mega" earthquakes in this region of Japan. The previous earthquake in this 500-year cycle was the 1611 Keicho earthquake, which also caused a devastating tsunami in Northeast Japan [5].

We propose to bring in a new type of seismic observation to aid the forecasting of earthquakes, namely by using travel times of seismic waves, which were recorded by seismic stations from local earthquakes in areas of large earthquakes. From travel times the inherent seismic velocity and other elastic properties of the Earth between the earthquake source and the recording station can be determined. In a number of studies information on velocity is applied for monitoring of stress field in areas of strong earthquakes (see e.g [11]). However the time dependence of the velocity parameter behavior before disaster is not much explored. In this chapter we formulate the research problem to study the behavior of this parameter, which can be reconstructed by using the travel times of P-waves and by applying tomography methods. Thus additional observations can be involved to solve the earthquake forecasting problem. It is assumed that 3-D velocity images can be analyzed during a long time period before the appearance of large event. The time window can be selected on the base of particularities of seismic activity in a region. Such 4-D imaging will allow us study the complex system of strong earthquakes.

2 A Method to Construct the 3-D Velocity Model for Each Time Window

The Differentiated Approach (DA) is a flexible recently developed method to resolve a complex velocity structure [14]. It is based on an algorithm called the Modification of Gaussian Elimination (MGE) which is a new numerical scheme [13]. The advantage of DA is its stability, in spite of data errors, that characterize all measurements (in the seismic case they include observation and modeling errors that lead to errors in hypocenter determination). Additional advantage is the ability of the MGE method to effectively solve large sparse linear systems with the possibility to illuminate structures in details almost down to their inherent physical measuring resolution, which is e.g. governed by the frequency range of the observed data (Fresnel zone radius). In other words, there is no numerical limitation in DA. With conventional methods the smaller the block-size of a modeled structure, the larger is a size of the initial system of equations, which is computationally intensive to solve.

With the MGE any small size of model block (cell) can be chosen, because MGE divides the initial system into a set of subsystems and thus estimates the resolution of smaller subsystems. An estimation of the resolving power of the least squares method for a large initial set of equations is a highly challenging problem, even for the best of modern computers.

In selecting a robust solution the DA method applies statistical evaluation of the solution. The method, however, requires relatively good coverage of the study area by seismic rays. The greater the number of seismic rays that intersect a target block from different directions, the higher the probability of a reliable estimate.

Let us consider an example of the DA method. Figure 1 shows the P-wave velocity image of the Tjörnes Fracture Zone (TFZ) in Northeast Iceland that has been

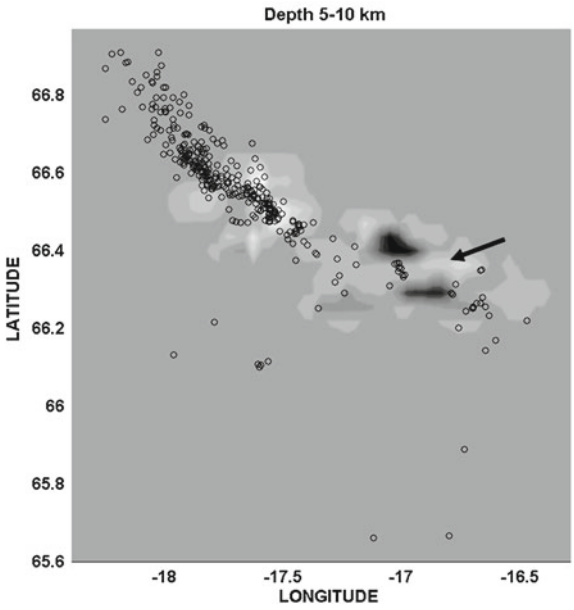


Fig. 1 P-wave velocity structure derived by DA for TFZ along the Grímsey lineament in Iceland. Low (high) velocity zones are colored in *white* (*black*). Hypocenters of local earthquakes are denoted by *open circles*. The *black arrow* shows neighboring low and high velocities that correspond to the greatest jump in the distribution function (see Sect. 3)

constructed with travel times from local earthquakes recorded between 1986 and 1988 [12]. Thus the length of the time-window was 3 years. In Sect. 3 we will analyze the given velocity model in detail.

3 4-D Velocity Tomography via Distribution Function

The construction of tomographic images in various regions of the world shows that 3-D velocity images of the same area are different for different time windows. It could be due to random character of seismic events. For a given time window, the 3-D image is determined by set of velocity values that were found in blocks of the geophysical model. These values are real numbers $v_1, v_2, \ldots, v_k$, which can be considered as the result of a random experiment. The vector $v = (v_1, v_2, \ldots, v_k)$ has a dimension of k, which is defined by the number of blocks of the medium and it is a random variable.

The probability of distribution of the random variable is characterized by the distribution function $F(x)$, which plays a fundamental role in mathematical statistics. According to the definition,

$$\begin{array}{l} F(x) = P(v \le x) \\ 0 \le F(x) \le 1 \end{array},$$

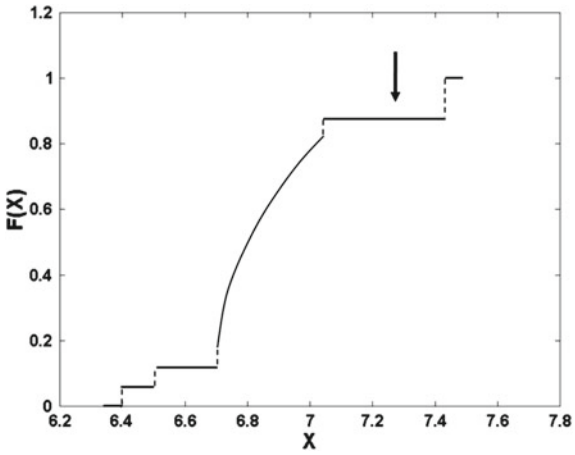


Fig. 2 The *horizontal axis* is the velocity random value. The *vertical axis* is the distribution function value. The *black arrow* shows the greatest jump of the distribution function

where P denotes a probability or the repetition frequency of an event $\{v \leq x\}$. $P = k1/k$, where $k1$ is a number of the repetitions of an event in a given random experiment.

We will apply this function in order to estimate the velocity field that determines a structure of the geophysical medium for a given time window. Jumps in the distribution function correspond to discontinuities in velocity values or contrasts in the velocity distribution. Figure 2 shows the distribution function that was constructed using velocity values, which correspond to the image in Fig. 1. The distribution function can be decomposed into a discontinuous (stepped) part and a continuous part that can be obtained by approximating discrete values via a curve. The greatest jump of the function is denoted by the arrow in Fig. 2.

Location of the largest velocity jump is indicated in Fig. 1. This jump corresponds to a low velocity zone (colored in white) that is sandwiched between two high velocity anomalies (colored in black). In total, the distribution function has 5 discontinuity points: x = 6.4; 6.5; 6.7; 7.0; 7.4. The dotted line shows the position of each point. From the point of view of statistics the probability mass is concentrated at discontinuity points [2]. This means that the listed velocity values characterize P-wave velocity with a high probability for the given time window along Grímsey lineament in TFZ, Northeast Iceland.

We have analyzed a behavior of the distribution function for a single sample, which corresponds to a fixed time window. The time behavior of this function can be determined, if it is inspected in multiple time windows, which together make up a longer period of time. The problem can be reduced to the construction of the sequence of distribution functions. We assume that a series of random experiments will allow us to get a stable result, which determines a behavior of the geophysical parameter (velocity of seismic waves) before the onset of strong earthquakes.

4 Conclusion

The seismic wave velocity of the crust is directly related to the state of the stress field [11] and it is an important parameter for the complex system of strong earthquakes. We assume that a change in the stress field before the occurrence of large earthquakes will influence this parameter, and that it may be possible to determine its critical values. Hence it is necessary to establish its critical values. In this chapter we propose to analyze a seismic velocity as a random variable and to study it via a distribution function. We will formulate the forecasting of strong earthquakes by means of constructing of series of 3-D velocity structures and distribution functions respectively in order to understand complex dynamics of a geophysical medium before the earthquake. Statistical analysis of distribution functions will reveal the stable characteristics of tomographic models, which may include time dependent features.

Acknowledgments We thank the DAAD foundation (Germany) for support, due of which tomography image that we used in this chapter has been constructed. Our thanks go to Prof. Wolfgang Jacoby (Mainz University, Germany) for fruitful discussion of tomography results.

References

1. Abe, S., Suzuki, N.: Main shocks and evolution of complex earthquake networks. Braz. J. Phys. **39**(2A), 428–430 (2009)
2. Cramer, H.: Mathematical Methods of Statistics. University of Stockholm (1946)
3. Hasegawa, A., Yoshida, E., Okada, T.: Nearly complete stress drop in the 2011 Mw 9.0 off the Pacific Coast of Tohoku Earthquake. Earth Planet. Space. **63**, 703–707 (2011)
4. Hirata, N., Yokoi, S., Nanjo, K.Z., Tsuruoka H.: A forecast experiment of earthquake activity in Japan under collaboratory for the study of earthquake predictability (CSEP). In: Geophysical Research Abstracts of EGU General Assembly. 14. EGU2012-4350-1. EGU, Vienna, Austria (2012)
5. Koketsu, K., Yokota, Y., Kato, N., Kato, T.: Identification and simulation of seismic supercycles along the Japan trench including the 2011 Tohoku Earthquake. In: Geophysical Research Abstracts of EGU General Assembly. 14. EGU2012-9357. EGU, Vienna, Austria (2012)
6. Lyapunov, A.M.: Works in 3 Volumes (in Russian). Publishing House of the USSR Academy of Sciences, Moscow-Leningrad (1954–1959)
7. Maccaferri, F., Rivalta, F., Passarelli, L., Jonsson, S.: The stress shadow induced by the 1975–1984 Krafla Rifting Event. In: Geophysical Research Abstracts of EGU General Assembly. 14. EGU2012-2584. EGU, Vienna, Austria (2012)
8. Nikolis, G., Prigogin, I.: Exploring Complexity, an Introduction. W.H. Freedman and Co., New York (1989)
9. Rebetskii, Y. L.: Estimation of the stress field values in the method of cataclastic analysis of shear fractures. Dokl. Earth Sci. **428**(7), 1202 (2009)
10. Rebetsky, YuL, Kuchai, O.A., Sycheva, N.A., Tatevossian, R.F.: Development of inversion methods on fault slip data stress state in orogenes of the Central Asia. Tectonophysics **581**, 114–131 (2012)
11. Slavina L.B., Myachkin V.V., Kuzmina T.A.: On the location and time of travel time precursors before large earthquakes. J. Earthq. Predict. Res. **2**(4), 375–384 (1993)
12. Smaglichenko, T., Jacoby, W., Fedorova, T., Wallner, H.: Stable estimate of velocity anomalies around Grimsey Lineament (Tjornes Fracture Zone, Iceland) with differentiated

tomography. In: Geophysical Research Abstracts. EGU General Assembly. 11, EGU2009-1332. EGU, Vienna, Austria (2009)
13. Smaglichenko, T. A.: Modification of Gaussian Elimination for the Complex System of Seismic Observations. Founded by Stephen Wolfram, vol. 20. Issue 3, pp. 229–241. Complex Systems Publications, Inc., USA (2012)
14. Smaglichenko, T. A., Shigeki, H., Kaori, T.: A differentiated approach to the seismic tomography problem: method, testing and application to the Western Nagano fault area (Japan). Int. J. Appl. Earth Obs. Geoinf. (Elsevier) **16**, 27–41 (2012)
15. Stefánsson R.: Advances in Earthquake Prediction: Seismic Research and Risk Mitigation, p. 264. Spinger-Verlag, Berlin (2011)
16. Zöller, G., Hainzl, S., Holschneider, M.: Recurrence of large earthquakes: Bayesian inference from catalogs in the presence of magnitude uncertainties. Pure Appl. Geophys. **167**(6–7), 845–853 (2012)

Tomography Application to Complex Seismic Data of the Tjornes Fracture Zone (Iceland)

Maria K. Sayankina, Tatyana A. Smaglichenko and Wolfgang R. Jacoby

Abstract The Tjörnes Fracture Zone (TFZ) is an active seismic zone in Northeast Iceland. It plays a key role to understand the geodynamic movement and location of tectonic plates. However the seismic experiment can not be performed close to earthquake sources, because sources are mainly located in the Greenland Sea. The unusual geological structure of TFZ and the limited conditions of an experiment lead to significant deviations between real observations and the values that are calculated in accordance with theoretical models. Consequently, there is a loss of adequacy and stability of tomography systems. Outcomes of the method, which takes into account this problem, are analyzed in the present chapter. Values and locations of P-wave velocity anomalies are comparing with those, which previously obtained from studies in geochemistry field and with GPS data.

Keywords Seismic data offshore · Tomography · Stability of linearized systems

1 Introduction

In North Atlantic the line of earthquakes is distributed along the Rekyanes Ridge. It cuts the central part of Iceland, and then continues their development along the Kolbesey Ridge. Seismic activity defines the geometry of the North American and

M. K. Sayankina (✉) · T. A. Smaglichenko
Research Oil and Gas Institute of Russian Academy of Sciences, Moscow, Russia
e-mail: msayankina@gmail.com

T. A. Smaglichenko
e-mail: t.a.smaglichenko@gmail.com

W. R. Jacoby
Institut für Geowissenschaft, Mainz University, Mainz, Germany
e-mail: jacoby@mail.uni-mainz.de

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_38,

Eurasian plates, the first of which is moving to the east and the other goes to the west with offset of 1–3 cm/year [8]. The plates move along the Tjornes Fault Zone (TFZ) [2], which is mainly located in the Greenland Sea between the Northeast Iceland and the Kolbeinsey Ridge. TFZ includes two major lineaments, the Husavik Fault Flatey (HFF), which is partly exposed on land and the Grimsey Lineament (GL) that is defined by offshore earthquakes.

Seismic tomography is often based on an assumption about the theoretical model that describes the change of a seismic velocity with a depth. Because of an offshore location the TFZ models (of about 10) are based on submarine profiles data (e.g. [4, 10]). These data show the TFZ structure complexity. The difference in velocity values obtained on the same depth is not a small. It becomes too difficult to find the appropriate 1-D model that can be used as the initial one in order to construct 3-D tomographic images. There is a wide range of differences between observed and calculated travel times and we can get a set of inversion solutions of the linearized system, some of which are nonrealistic values.

In this paper we consider the result of the tomographic approach, according to which seismic data are selected to construct the solution that is maximally free of data errors influence. The base of the approach is a numerical scheme that is called the Modification of Gaussian Elimination (MGE) [12]. In Sect. 3 the method is briefly described. Then we investigate how the result obtained with this method is consistent with outcomes of geophysical and petrologic studies.

2 Distribution of Stations and Sources and Wide Range of Observation Data

Our data set consisted of 574 earthquakes having small magnitudes. Events were recorded by 18 portable seismic stations of Mainz University, Germany during 1986–1989. Figure 1 shows that quakes were basically distributed as clusters along GL in the Greenland Sea while stations were mainly located in the coastline along the Tjornes peninsula.

Assumptions about the 1-D initial velocity model were made on the basis of submarine observations conducted in the work [14]. Namely 18 ocean bottom seismometers were installed along profiles and 35 explosions were carried out in the Greenland Sea. Authors have processed records by applying seismic methods. Values of seismic velocities were obtained in local points of TFZ. The theoretical 1-D model was constructed by using these values. Travel times were calculated using the found model. Statistical characteristics of a vector of differences between calculated and observed values were the next. Mean value (m) and standard deviation (std) were equal to -0.18 and 0.23, respectively. The number of negative values was much bigger than the number of positive ones. This tells us about asymmetry of observation data and low probability of closeness between theoretical and real values. By using parameters of the initial model and applying a search algorithm to construct the

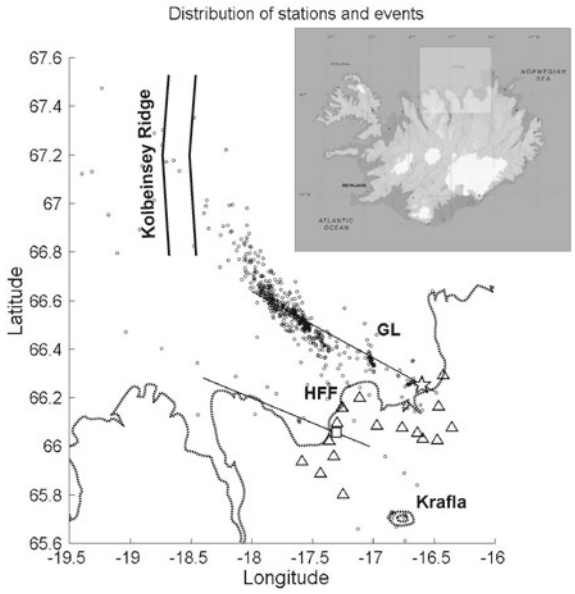


Fig. 1 The Tjornes Fracture Zone. *HFF* and *GL* are the Husavik Fault Flatey and the Grimsey Lineament *Open triangles* and *dots* indicate seismic stations and earthquakes, respectively. The TFZ location is shown by *square* on Iceland map in the right corner

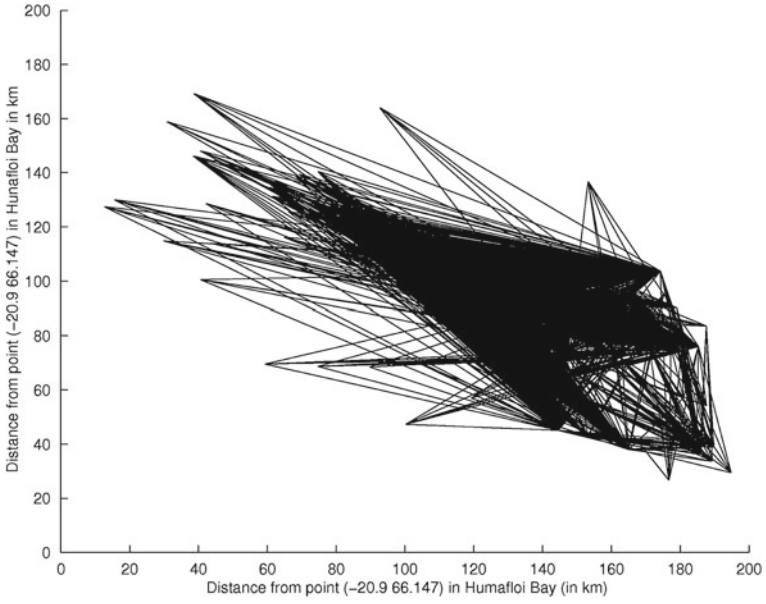


Fig. 2 The XOY projection of seismic rays in the Cartesian coordinate system that was turned so that the horizontal axis was parallel to the HFF location

minimal 1-D model [11] we got symmetrical distribution of observation data with m = 0.028 and std = 0.23. Figure 2 shows the XOY projection of seismic rays, lengths of which were calculated in accordance with the minimal model.

3 A Search for the Stable Solution of the Linearized System

Tomographic problem is reduced to the solution system of linear equations:

$$Au = \delta t + \varepsilon \tag{1}$$

where A is the matrix of lengths of seismic rays in blocks (cells of a medium), u is the unknown vector of velocity perturbations in blocks, δt is the known vector of a difference between observed and calculated travel times, ε is the vector of errors that involve modeling, measurement errors and errors in location of sources. The system is over-determined, because in seismic experiment the number of seismic rays is always greater than the number of blocks.

Large variations of values in the right-hand side of a system generate unrealistic values of solution. Seismic rays passing through the same blocks and having different observation values cause the system instability. To find stable values of a solution the MGE method has been used. The system of (1) was divided into small sub-systems so that components of an unknown vector were the same in few sub-systems. For each sub-system the basic matrix was found in accordance with criteria that are described in the work [13]. Basic solutions for the same unknown component were compared. The stable value of solution was selected for each component.

4 A Comparison of Tomographic Results with Other Studies of TFZ

Let us consider the P-wave velocity image that was obtained with the MGE method in the upper part of the crust in the depth range 0–5 km. In Iceland active volcanic areas are connected with hydrothermal activity. Hydrothermal systems can be both terrestrial and submarine. Hydrothermal fields are located in sediment layers. The area we study in this chapter includes one of the hydrothermal fields. This field is located in the northern part of TFZ and it is called by the Grimsey hydrothermal field.

A careful analysis of geochemical samples of rocks collected in the central and northern parts of the Grimsey field at a depth of about 400 m was carried out in laboratories of Freiberg University [7]. Hydrothermal mounds are mainly composed of hard anhydrite, gypsum, hydrothermal clay. It was also found that the width of mounds reaches 1 km. Figure 3 shows a location of the Grimsey field on the map of velocity anomalies that were determined with MGE. The seismic velocity that was found for the block that includes the field is equal to 4.89 km/s. This value is higher than the initial velocity, which we assumed for the depth range of 0–5 km. In the north end of GL one can see different patterns in the velocity distribution. The low-velocity anomaly (colored in a white) and the high-velocity anomaly (colored in a black) are neighbors. This jump in seismic velocities characterizes the sharp

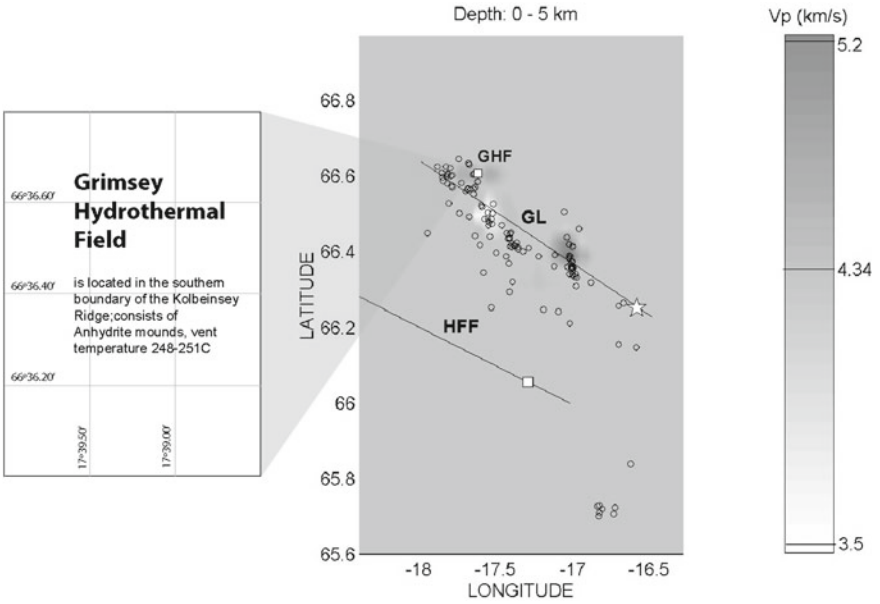


Fig. 3 P-wave velocity structure that found due to MGE at the depth range 0–5 km. *Open circles* denote hypocenters of earthquakes. *HFF*, *GL* and *GHF* are Husavik Fault Flatey, Grimsey Lineament and Grimsey Hydrothermal Field. On the *HFF* line the *white rectangle* corresponds to a location of the Husavik village. The epicenter projection of the 1976 Kópasker earthquake is denoted with the *white star*

change in the density of rocks. We suggest that the density change corresponds to the presence of the hydrothermal field within the high velocity zone. The found numerical value of velocity in the Grimsey field is in a good agreement with the value of P-wave velocity, which was defined in the fault zone that is suited in the Coso geothermal field, California [9]. Authors estimated P-wave velocity of 4.8 km/s in this zone by comparing the observed waveform with the synthetic calculated ones.

Note that the hypothesis about hydrocarbon deposits in hydrothermal areas of the Mid-Atlantic Ridge was explained and suggested in the work [3]. This hypothesis was confirmed by first observations in the Grimsey hydrothermal field [5]. It was found that the predominant gases are hydrocarbons [1].

As told above the seismic activity of Iceland divides land into parts. It is linked to the intensity of volcanic activity, including the Krafla caldera. This causes a displacement of the crust. The models of displacement fields were built by authors of the work [6]. By comparison of simulation results and GPS observations the location of fields has been determined with a small error. We compared the location of a displacement field that is connected with the Krafla volcano with velocity structure, we have found in the depth range of 5–10 km. Figure 4 shows that this field extends beyond the coast line at the north of Iceland, and corresponds to the low velocity anomaly that we found near projection of hypocenter of the damaging 1976 Kopasker earthquake

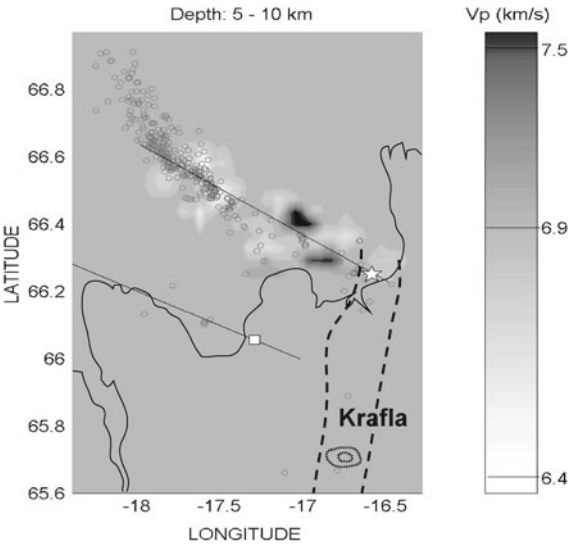


Fig. 4 P-wave velocity structure that found due to MGE at the depth range 0–5 km. On the *HFF* line the *white rectangle* corresponds to a location of the Husavik village. The epicenter projection of the 1976 Kópasker earthquake is denoted by the *white star*. The *dotted line* corresponds to boundaries of the displacement field that is connected with the Krafla caldera

(indicated by a white star). According to the study [6] the displacement field locked at a depth of 10 km.

We assume that low velocities are in an agreement with the location of displacement field that is a zone of high seismic risk.

5 Conclusion

We have considered the application of the most recent developed method to the TFZ seismic data, the complexity of which is caused by uncomfortable natural conditions to run an experiment. Moreover the data set is a small: about 4,000 seismic rays. The number of rays should be much more if we use the conventional tomography method [10]. In spite of these facts the outcome of the MGE method fairly realistic, because it is consistent with the results of other observations including chemical analysis and GPS data. Notice that the cost we "paid" for the result is a half of initial data set. The method selects the rays that give consistent results and reject others. On this basis we can conclude that to get a reliable solution we need to do as much as possible seismic observations and measurements.

Acknowledgments We are very grateful to the DAAD foundation for support of research of the innovative method, result of which is partly presented in this chapter.

References

1. Atkins, D.: Exploration techniques for locating offshore geothermal resources. Master of Science in Sustainable Energy Thesis, Reykjavik University (2013)
2. De Mets, C., Gordon, R.G., Argus, D.F.: Geologically current plate motions. Geophys. J. Int. **181**(1), 1–80 (2010)
3. Dmitrievsky, A.N., Karakin, A.V., Balanyuk, I.E., Matveenkov, V.V.: Hydrothermal mechanism of hydrocarbons in the mid-ocean ridges (on examples of the Barents and Norwegian Seas). Oil Gas Geology. **8**, 4–16 (1997). (in Russian)
4. Flovenz, O.G., Gunnarson, K.: Seismic crustal structure in Iceland and surrounding area. Tectonophysics **189**, 1–17 (1991)
5. Hannington, M., Herzig, P., Stoffers, P., Scholten, J., Botz, R., Garbe-Schonberg, D., Jonasson, I.R., Roest, W.: Shipboard scientific party: first observations of high temperature submarine hydrothermal vent and massive anhydrite deposits off the North Coast of Iceland. Mar. Geology. **177**, 199–220 (2001)
6. Jouanne, F., Villemin, T., Ferber, V., Maveyraud, C., Ammann, J., Henriot, O., Got, J.-L.: Seismic risk at the rift transform junction in North Iceland. Geophys. Res. Lett. **26**(24), 3689–3692 (1999)
7. Kuhn, T., Herzig, P.M., Hannington, M.D., Garbe-Schönberg, D., Stoffers, P.: Origin of fluids and anhydrite precipitation in the sediment-hosted Grimsey hydrothermal field North of Iceland. Chem. Geol. **202**(1–2), 5–21 (2003)
8. Lawver, L., Müller, R.: Iceland hotspot track. Geology **22**, 311–314 (1994)
9. Lou, M., Malin, P. E., Rial, J. A.: Locating an active fault zone in Coso geo-thermal field by analyzing seismic guided waves from microearthquake data. In: Proceedings of 20th Workshop on Geothermal Reservoir Engineering. Stanford Universlty, California, pp. 115–121 (1995)
10. Riedel, C., Tryggvason, A., Dahm, T., Stefanson, R., Bodvarson, R., Gud-mundsson, G.B.: The seismic velocity structure north of Iceland from joint inversion of local earthquake data. J. Seismolog. **9**, 383–404 (2005)
11. Smaglichenko, T., Jacoby, W., Wallner H.: Tomographic P-velocity crustal images for the Tjornes fracture zone (NE Iceland): a new minimum 1-D model. In: News Letter of European Geophysical Society, Nice, France, vol. 78, p. 51 (2001)
12. Smaglichenko, T.A.: Modification of Gaussian Elimination for the Complex System of Seismic Observations. Founded by Stephen Wolfram, vol. 20. Issue 3, pp. 229–241. Complex systems Publications, Inc., USA (2012). http://www.complex-systems.com/index.html
13. Smaglichenko, T. A., Shigeki, H., Kaori, T.: A differentiated approach to the seismic tomography problem: method, testing and application to the Western Nagano fault area (Japan). Int. J. Appl. Earth Obs. Geoinf. (Elsevier) **16**, 27–41 (2012)
14. Sturkell, E., Brandsdottir, B., Shimamura, H., Mochizuki, M.: Seismic crustal structure along the Axarfjordur Trough at the eastern margin of the Tjornes fracture zone. N-Iceland. Jokull. **42**, 13–23 (1992)

A Complexity of the Displacement Along Segments of the Akhtyirskiy Fault

Alexander V. Smaglichenko, Lidia A. Sim and Andrey V. Gorbatikov

Abstract A distinctive features of the Ahtyirskiy fault (Krasnodar, Russia) are its long length and the unusual geological structure. The fault demarcates the boundary between highlands and flat landform. To study the fault displacement we have applied two methods. The first structural–geo-morphological method reconstructs orientation of the compression/expansion axis in a horizontal plane and thus determines the direction of a horizontal displacement along the fault. In this study the method has been modified by analyzing of individual segments of the Ahtyirskiy fault. Another method is the microseismic sounding that determines a distribution of surface wave characteristics in a vertical plane. We conclude that the displacement along the fault has average direction, which we call the upthrust right-shift. However there are segments that are deformed in a special manner and displacements along them can not be explained via theoretical models. Mainly the fault plane falls to the south direction while its individual parts fall to the north.

Keywords The fault displacement · Structural geology · Seismic methods

1 Introduction

The Ahtyirskiy fault line runs between different geological structures. The fold structure is located at the southern end while the Predkubanskiy bending is settled at the north. The fault zone is situated in the Black Sea–Caucasus region, which

A. V. Smaglichenko (✉) · L. A. Sim · A. V. Gorbatikov
Schmidt Institute of Physics of the Earth of Russian Academy of Sciences, Moscow, Russia
e-mail: asmaglichenko@dev.rtsoft.ru

L. A. Sim
e-mail: sim@ifz.ru

A. V. Gorbatikov
e-mail: avgor70@mail.ru

A. Sanayei et al. (eds.), *ISCS 2013: Interdisciplinary Symposium on Complex Systems*, Emergence, Complexity and Computation 8, DOI: 10.1007/978-3-642-45438-7_39,

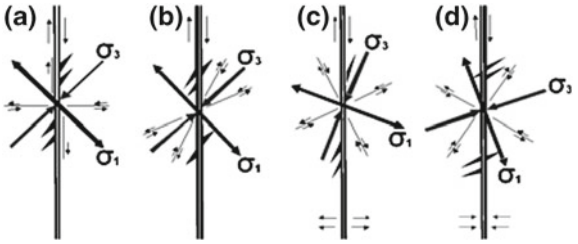


Fig. 1 The Gzovsky' model of the right-shift displacement (from *left* to *right*): **a** a normal right shift on 45° **b** a right shift at a sharp angle **c** a right shift with the further extension **d** a presence of an additional compression

is characterized by a weak seismic activity. The particular feature of the Ahtyirskiy fault is that it has a lot of branches that are visible at the surface.

There are several points of view regarding the origin of fault branches. And the question about the horizontal component of a fault displacement is the hot topic for discussion. For example, there is an opinion [3] that this component has the right-shift character. In this work we investigate the fault by considering it as the segments system and by analyzing a displacement of each segment.

To get extended knowledge about the fault and its segments we apply the structural–geo-morphological method [4, 5] by using the satellite photographic images that were obtained via Google Earth Pro. To understand the process of vertical movements of particular segments we analyze the seismic model, which was constructed by the micro-seismic sounding method on the base of a seismic recording of surface waves [2].

2 The Character of Displacements Along the Akhtyirskiy Fault

Satellite images detect various features of the geological structure on the Earth's surface. Relief lines on images are the result of a dynamic motion of the Earth's crust, which is always under influence of stress fields especially in fault zones. Movements occur in different planes and directions depending on the orientation of tectonic stresses, which form the geometry of the fault and its activity.

Displacements along the fault are accompanied by the formation of joints [6]. In photographic images the joints correspond to small rectilinear elements of a relief. We will call these elements by mega-joints. According to the structural-geo-morphological method the orientation of mega-joints with respect to the fault line indicates the direction of a horizontal displacement and characterizes the stress field. The mutual orientation of the fault line and gaps in the distribution of joints were modeled by Gzovsky [1]. By analyzing of the distribution of mega-joints that are visible on images we study the correlation between them and theoretical models, which were developed by Gzovsky. In the case of the Akhtyrsky fault we determined that the most distributed mega-joints are in agreement with the model of right-shift displacements (Fig. 1).

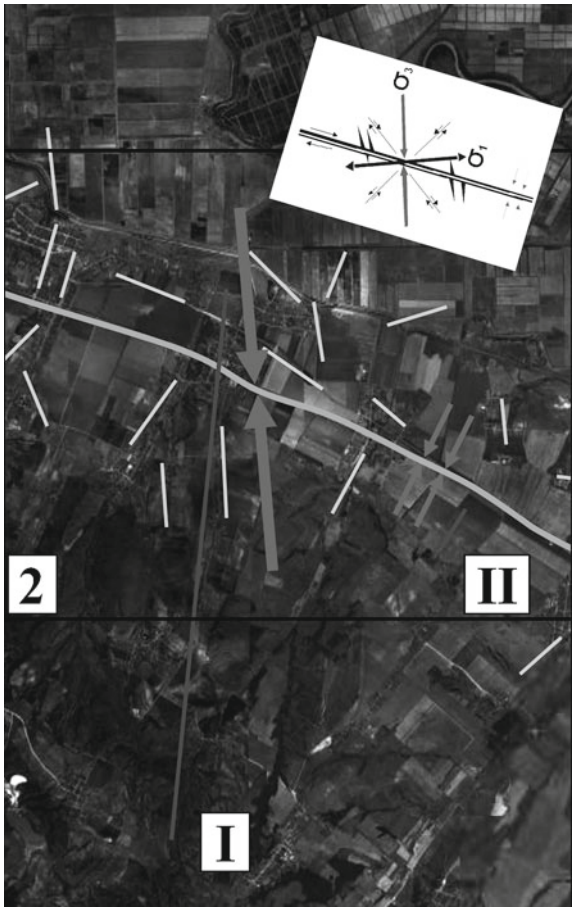


Fig. 2 The reconstructed field of displacements along the segment with number 2 Profiles are indicated by symbols *I* and *II*. The variant of the Gzovsky' model (on the *corner*) is in good agreement with the distribution of mega-joints

At the same time the satellite image showed that not all mega joints had orientations in accordance with the theoretical model. Therefore the fault area was divided into five segments. Individual segments having the best conformity of the theoretical model were selected. Figure 2 shows an example of such segment.

Nevertheless we have found that there are segments, which can be simultaneously explained via two variants of the model. This tells us about the complexity of the fault. Figure 3 demonstrates this example.

We assume that such situation can be explained by the further expansion of the fault in its north-western end. This can be related to the formation of cavities in the Black and Azov Seas. However we also may suggest that such character of displacements can be connected with inhomogeneous structure beneath the fault plane.

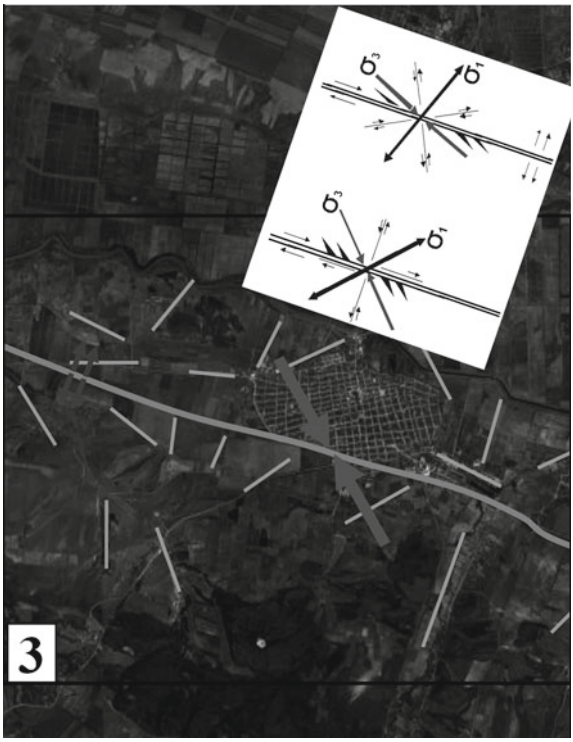


Fig. 3 The reconstructed field of displacements along the segment with number *3*. Two variants of the Gzovsky' model (on the *corner*) correspond to the mega joints distribution

3 Vertical Components of Displacements Along the Akhtyirskiy Fault

The method of micro-seismic sounding uses observation data of a seismic noise. There are many causes of a noise. It can be created due to the nature and also owing to human activity. For example, if seismic measurements are performed near a big plant, then the seismic records of surface waves will detect vibrations, which are basically created by the work of this plant. The method of micro-seismic sounding uses natural micro-seismics of the Earth. The technology is based on the inversion of micro-seismic field parameters [2].

Two images have been constructed applying this method. They show the change of an intensity of a surface wave with depth. Seismic experiments were conducted along two profiles (see Fig. 2). The vertical section along the profile I (Fig. 4) shows complicated structure of the fault zone. The boundary of the intensity anomaly (colored in a black) falls up to the depth of 10 km in the upper crust to the south direction. On the whole the intensity distribution is in an agreement with hypothesis about the up-thrust component of the fault displacement.

The second profile (denoted by II in Fig. 2) is much shorter than the first one. The results obtained for this profile show the presence of the vertically distributed

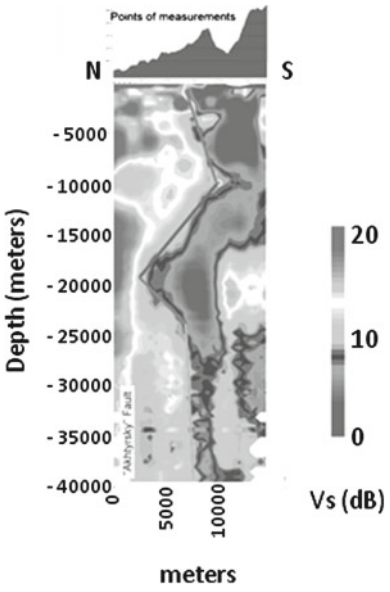


Fig. 4 Surface wave intensity distribution along the first profile (I). The anomaly that has maximal value of the intensity of surface wave is colored in a *black*

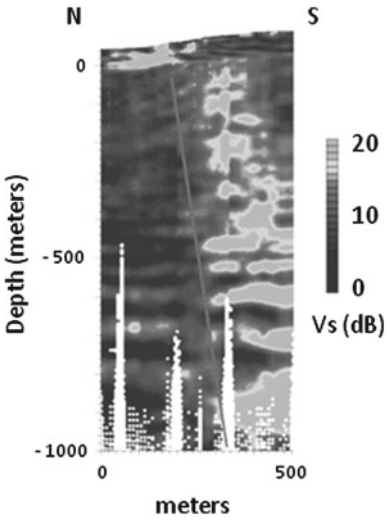


Fig. 5 Surface wave intensity distribution along the second profile (II). The surface wave intensity anomaly is colored in a *light gray*

anomaly till the depth of 1 km (Fig. 5). By analyzing the relief around the place, where the experiment was conducted one can assume that the anomaly is connected with some branch of the fault, which is moving to a surface.

Thus the method of micro-seismic sounding reveals the presence of vertical displacents while the structural–geo-morphological method defines the horizontal displacement. Hence combination of two methods extends our knowledge about the fault movement.

4 Conclusion

Based on this study, we conclude that the application of methods that correspond to two different disciplines of geology and seismology leads to the better understanding of the character of displacements along the Akhtyrskaya fault. The complexity of the fault movement can be determined by analyzing the fault as a chain of its individual segments. Owing to such approach we detected the presence not only horizontal but also vertical displacements in the fault zone.

Acknowledgments We thank Yuri Rebetskiy, Anastasia Mikhaylova and other colleagues of the Gzovsky's laboratory of Shmidt's Institute of Physics of the Earth for fruitful discussion of results that are presented in this chapter.

References

1. Gzovsky, M. V.: Mathematics and Geotectonics. Nedra (1971) (in Russian)
2. Gorbatikov, A.V., Stepanova, MYu.: Microseismic field affected by local geological heterogeneities and microseismic sounding of the medium. Phys. Solid Earth **44**(17), 577–592 (2008)
3. Rastsvetaev, L. M.: Tektono-dynamic conditions of formation of the structure of large-alpine of the Caucasus (in Russian). Geology and Mineral Re-sources of the Great Caucasus, pp. 69–96. Nauka (1987)
4. Sim, L.A.: The study of tectonic stress by using the geological indicators (methods, results, recommendations) (in Russian). Geol. Explor. **10**, 3–27 (1991)
5. Sim, L.A., Sergeev, A.A.: Eine Strukturell—Geomorphologische Methode zur Analyze Aktiver Bruchemit dem Zeit der Bestimmung Neotektonischer Spannungen in Tafelgebieten. Ztschr. Geol. Wiss. **20**, 369–375 (1996)
6. Sim, L.A., Mikhaylova, A.V.: Faults, the sedimentary covers, platforms and methods of their study (in Russian). Prob. Tectonophys. 141–148 (2008)