

Furtado, Bernardo Alves; Sakowski, Patrícia Alessandra Morita

Working Paper

Complexity: A review of the classics

Texto para Discussão, No. 2019a

Provided in Cooperation with:

Institute of Applied Economic Research (ipea), Brasília

Suggested Citation: Furtado, Bernardo Alves; Sakowski, Patrícia Alessandra Morita (2014) : Complexity: A review of the classics, Texto para Discussão, No. 2019a, Instituto de Pesquisa Econômica Aplicada (IPEA), Brasília

This Version is available at:

<https://hdl.handle.net/10419/121509>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

2019a

DISCUSSION PAPER

COMPLEXITY: A REVIEW OF THE CLASSICS

Bernardo Alves Furtado
Patrícia Alessandra Morita Sakowski



2019a

DISCUSSION PAPER

Brasília, December 2014

COMPLEXITY: A REVIEW OF THE CLASSICS¹

Bernardo Alves Furtado²

Patrícia Alessandra Morita Sakowski³

1. The authors appreciate the reading and comments of Leonardo Monasterio, Nilo Saccaro Jr. and Júlio Cesar Roma. Errors and omissions remain our own.

2. Researcher at the Department of Regional, Urban and Environmental Studies and Policy at Ipea and CNPq researcher.

3. Researcher at the Department of Regional, Urban and Environmental Studies and Policy at Ipea.

Federal Government of Brazil

**Secretariat of Strategic Affairs of the
Presidency of the Republic**
Minister Marcelo Côrtes Neri

ipea Institute for Applied
Economic Research

A public foundation affiliated to the Secretariat of Strategic Affairs of the Presidency of the Republic, Ipea provides technical and institutional support to government actions – enabling the formulation of numerous public policies and programs for Brazilian development – and makes research and studies conducted by its staff available to society.

President

Sergei Suarez Dillon Soares

Director of Institutional Development

Luiz Cezar Loureiro de Azevedo

Director of Studies and Policies of the State, Institutions and Democracy

Daniel Ricardo de Castro Cerqueira

Director of Macroeconomic Studies and Policies

Cláudio Hamilton Matos dos Santos

Director of Regional, Urban and Environmental Studies and Policies

Rogério Boueri Miranda

Director of Sectoral Studies and Policies, Innovation, Regulation and Infrastructure

Fernanda De Negri

Director of Social Studies and Policies, Deputy

Carlos Henrique Leite Corseuil

Director of International Studies, Political and Economic Relations

Renato Coelho Baumann das Neves

Chief of Staff

Bernardo Abreu de Medeiros

Chief Press and Communications Officer

João Cláudio Garcia Rodrigues Lima

URL: <http://www.ipea.gov.br>

Ombudsman: <http://www.ipea.gov.br/ouvidoria>

DISCUSSION PAPER

A publication to disseminate the findings of research directly or indirectly conducted by the Institute for Applied Economic Research (Ipea). Due to their relevance, they provide information to specialists and encourage contributions.

© Institute for Applied Economic Research – **ipea** 2014

Discussion paper / Institute for Applied Economic

Research.- Brasília : Rio de Janeiro : Ipea, 1990-

ISSN 1415-4765

1. Brazil. 2. Economic Aspects. 3. Social Aspects.
I. Institute for Applied Economic Research.

CDD 330.908

The authors are exclusively and entirely responsible for the opinions expressed in this volume. These do not necessarily reflect the views of the Institute for Applied Economic Research or of the Secretariat of Strategic Affairs of the Presidency of the Republic.

Reproduction of this text and the data it contains is allowed as long as the source is cited. Reproductions for commercial purposes are prohibited.

JEL: A31, B15, C63

CONTENTS

ABSTRACT

SINOPSE

1 INTRODUCTION	7
2 MODELS	11
3 PATTERNS, DYNAMIC SYSTEMS, NON LINEARITIES AND CHAOS	14
4 CELLULAR AUTOMATA AND ARTIFICIAL INTELLIGENCE	19
5 INFORMATION THEORY	29
6 NETWORKS.....	36
7 EVOLUTION	39
8 CRITICISMS.....	42
9 CONCLUDING REMARKS	45
REFERENCES.....	47

ABSTRACT

This text was written as part of the project Modelling of Complex Systems for Public Policy. It reviews the classical authors who jointly contributed to establish the elements of what could constitute a “science of complexity”. Based on the original writings of these authors, the text discusses the central concepts of complex systems: *i)* the interaction between (homogeneous or heterogeneous) agents and the environment; *ii)* emergence and self-organization; *iii)* the importance of nonlinearity and scales; *iv)* the determinism of rules; *v)* the emphasis on dynamics and feedback; and *vi)* the notions of adaptation, learning and evolution. Finally, contemporary criticisms are presented. They suggest that the arguments of complex systems do not support the epistemological establishment of a supposedly new science, but they do not reject the advances proposed by complexity studies.

Keywords: complexity; emergence; dynamical systems; non-linearity; cellular automata; modelling; information theory; neural networks; evolution.

SINOPSE

Este texto está inserido no projeto Modelagem de Sistemas Complexos para Políticas Públicas e faz uma resenha dos autores clássicos que, em conjunto, contribuíram com os elementos do que seria uma “ciência da complexidade”. Com base no pensamento original destes autores, os conceitos centrais de sistemas complexos são discutidos, a saber: *i)* a interação entre agentes (homogêneos ou heterogêneos) e o ambiente; *ii)* as propriedades emergentes e a auto-organização; *iii)* a importância da não linearidade e das escalas; *iv)* as regras e seu determinismo; *v)* a ênfase na dinâmica e retroalimentação; e *vi)* as noções de adaptação, aprendizado e evolução. Por fim, críticas contemporâneas são apresentadas. Elas sugerem que os argumentos de sistemas complexos não sustentam epistemologicamente a constituição de suposta nova ciência, mas não rejeitam os avanços propostos nos estudos de complexidade.

Palavras-chave: complexidade; emergência; sistemas dinâmicos; não linearidade; autômatos celulares; modelagem; teoria da informação; redes neurais; evolução.

1 INTRODUCTION

Complex systems can be defined as systems in which agents interact with the environment and among themselves, with no central authority in a complex, non-linear and dynamic manner and whose actions generate emergent, observable structures through different scales from usually simple rules that together allow for adaptation and evolution of the system itself. Complex systems can be formally described and allow for computational simulation.¹

Melanie Mitchell proposes the following definition:

Systems in which organized behavior arises without an internal or external controller or leader are sometimes called self-organizing. Since simple rules produce complex behavior in hard-to-predict ways, the macroscopic behavior of such systems is sometimes called emergent. *Here is an alternative definition of a complex system: a system that exhibits nontrivial emergent and self-organizing behaviors.* The central question of the sciences of complexity is how this emergent self-organized behavior comes about (Mitchell, 2011, p. 13).

Simulating these systems from the public policy perspective means to model the action and the interaction among citizens, firms, institutions, and the environment constrained by the existing regulation, budgetary limits, policies' framework and spatial conditions. Thus, working with complex systems applied to public policy means creating experimental environments where the central characteristics of these systems are formally described and from which insights into policy can be obtained (relatively) easily and with low costs.

The objective of this text – which is part of the encompassing project titled Modeling Complex Systems for Public Policies – is to review the classic authors who have contributed to the constituent elements of a tentative “science of complexity”. The authors included in this review are based on the background literature put forward for the Complex Systems Summer School of 2014 by the Santa Fe Institute.²

1. The formal description of a complex problem is relatively simple. Miller and Page (2007, p. 11), for example, use the standard ovation problem to illustrate how a problem with difficult internal mechanisms can be formally described.

2. Available at: <<http://goo.gl/gYIW6l>>. For a good introductory book on the topic, see Mitchell (2011). Finally, see also the online course offered by the Santa Fe Institute at: <<http://goo.gl/e8FND4>>.

Given the proposed nature of this text as a panorama of reviews, the authors limit their analyses to their own interpretation of the classics' original works, searching only for a minimal chaining of concepts and broad themes.

The central concepts of complexity “emerge” within the context of research of each of the authors writing at different moments in time, and follow the typical jargon of each discipline. There is, therefore, superposition, addition and complementation at different levels of detail and specificity. Anyway, the definition of complex systems presented above aims at capturing all these elements.

In a rough and non-exhaustive way, we can summarize the contributions of the original authors as follows. Anderson (1972), Langton (1986) and von Neumann (1986) highlight the importance of *interaction among parts*. They defend that the lack of understanding of the connections across different scales compromises the understanding of the phenomenon. Simon (1973) confirms this view as he reinforces the value of hierarchies. Observed complexity, however, cannot demand that models equal reality, as is emphasized by Rosenblueth and Wiener (1945). They state that one should describe the minimum (necessary), the essence of the phenomenon, so that modeling can occur. Kolmogorov (1965), Shannon (1948), Chaitin (1966) and Gell-Mann and Lloyd (2004) all focus on how to measure this essence. Shannon (1948) defines the fundamental concepts of *information* such as the notions of noise, redundancy, and perturbation; source, transmitter, channel, and receiver, many of which seem automatic nowadays. Gell-Mann and Lloyd (2004) once again reinforce the relevance of scales and the basic tradeoff that scientists face, i.e., whether to gain in precision (and abstract the whole), or to generalize and understand relations at higher orders. Even considering the observed complexity of phenomena, it is possible to *distinguish patterns*, and behavioral classes, as established by Turing (1952). Wolfram (1984) confirms the patterns of Turing and realizes that systems *self-organize*. Langton (1986), in turn, demonstrates that those self-organized patterns occur highly frequently in real cases *in nature* (and in society). This self-organization leads to the *emergence* of phenomena, as Anderson (1972), Turing (1952), von Neumann (1966) and Hopfield (1982) had noted. Minsky (1961) and Turing (1950) explore the concepts of self-organizing, adaptation and modeling based on studies of learning and evolution in the fields of *computation and artificial intelligence*, in tune with the *information theory* of Shannon (1948). Waddington (1942) and other authors in the area of evolution emphasize the relevance of *adaptation*, and of the dynamic changes that these self-organizing systems go through.

Thus, the main features/aspects of complex systems used in the definition already given, namely:

- the interaction among agents (homogeneous or heterogeneous) and the environment;
- the emergent properties, the self-organization;
- the relevance of non-linearity, of the scale and of the hierarchies;
- the emphasis on the dynamics and feedback; and
- the notions of adaptation, learning and evolution.

Are all defined and detailed throughout the classic texts.

It is our hope to provide the reader with direct access to the original (and complex) thinking of the original authors, which as of now are available dispersedly, with some temporal linkage in some disciplines but not as yet consolidated as a single piece.

This paper is organized as follows. This introduction defines the general concept and establishes an initial epistemological approach (section 1); the next section (2) discusses models and regulation and section three describes dynamic systems, their rules and emergent behavior. Section 4 introduces cellular automata and its importance to computer science and artificial intelligence. The following section lays the basis of information theory (section 5) and section 6 provides notions of networks. Finally, brief concepts on evolution are presented (section 7). The paper concludes with “contemporary” criticisms on the “science of complexity” (section 8) and final considerations (section 9).

1.1 Epistemology

This section introduces the complementary concepts by Anderson (1972) and Simon (1973) on the relevance of including varied scales and the cross effects among scales when trying to understand complex, intricate phenomena.

Anderson (1972) starts his paper *More is different* disputing the hypothesis generally accepted by the scientific community of the reductionist paradigm – according to which the whole can be understood by the segmentation and understanding of each part individually.

Against reductionism the author argues that even if all natural laws were known, yet it would not be possible to rebuild the universe from the laws alone.

The behavior of large and complex aggregates of elementary particles, it turns out, is not to be understood in terms of a simple extrapolation of the properties of a few particles (Anderson, 1972, p. 393).

On the contrary, defends Anderson, at each level of complexity, new properties – which demand new understanding, new research – come up. The description of the form of an atomic nucleus, for instance, *depends on the scale of the analysis*. When the nucleus is considered to be a large *many-body system*, the description makes sense, but not otherwise. In order to illustrate the importance of scale, the author works on a molecular biology example to conclude that symmetry is central to physics. Yet, even in a symmetrical system, its internal structure may be asymmetric.

Still on the scale and large systems topic, Anderson emphasizes the nonlinearities that are present in the system:

[...in those systems] matter will undergo mathematically sharp, singular “phase transitions” to states in which the microscopic symmetries, and even the microscopic equations of motion, are in a sense violated (Anderson, 1972, p. 395).

Closing the argument the author states that “In this case we can see how the *whole becomes not only more than but very different from the sum of its parts*” (Anderson, 1972, p. 395, our emphasis). This quote has become a fundamental part of complexity as it is seen presently.

Finally, in order to conclude his paper, Anderson advocates the interdisciplinary approach – such as the complementarity that should exist between physics and chemistry – and quotes a dialogue between two authors in Paris:

- Fitzgerald: The rich are different from us.
- Hemingway: Yes, they have more money.

Whereas Fitzgerald focuses on the structural differences, Hemingway describes it as a quantitative element only.

Simon (1973) starts the paper establishing it as a *pion*,³ i.e., as the elementary particle that holds together the constituent elements of physicists and biologists. In fact, Simon suggests that “everything is connected, but some things are more connected than others” (Simon, 1973, p. 23). That also resembles Tobler’s first law of geography according to which “everything is related to everything else, but near things are more related than distant things” (Tobler, 1970, p. 236).

In the text, Simon also refers to boxes, in the sense used by Rosenblueth and Wiener (1945) when they describe recursive structures. The idea is that the boxes would be inside other boxes, which, in turn, would be inside other boxes, thus establishing complex hierarchies so that the higher the number of layers, the more complex the system. Besides, Simon defends the notion that the interaction among components of each box is more intense internally – within the same box – than externally across boxes.

Finally, Simon (1973) states his scientific preference towards the building up of *laws that are able to describe relations across hierarchical levels*, minimizing traditional approaches that simply describe the fundamental parts of the system.

In sum, the review of the authors that we propose aims at opening up windows of analysis and providing elements so that new interrelations can be made between the original concepts put forward and the readers’ background.

2 MODELS

Rosenblueth and Wiener (1945) departing point is the premise that science follows a dual rite that goes from the understanding of reality to the action upon it. The authors say that if one does intend to act upon the scientific phenomenon, thus exerting control, it is necessary to abstract reality, work with similar elements which are simpler, but which contain the essence.

The scientist – when at his or her working practice – does not have full methodological knowledge of their object of investigation. However, according to Rosenblueth

3. “A pion (short for pi meson, denoted with π) is any of three subatomic particles: π^0 , π^+ , and π^- . Each pion consists of a quark and an antiquark and is therefore a meson”. Available at: <<http://goo.gl/eUXxNp>>.

and Wiener, they should have a clear view about what their research questions are. These research question could be defined hierarchically from the most general to the most specific. The general question possibly contains more relevant answers, but they are also less precise and, especially, less susceptible to experimentation. From that, one can assume that the theoretical construction demands that the *variation on the scale* of the problem be more or less abstract. Besides, it also demands from the scientists the ability to translate from abstraction into experiments and vice versa.⁴

When defining models, the authors distinguish between material and formal models. Material models are representations of a system by another system which is simpler but possesses similar properties.⁵ Formal models are based on symbols, based upon ideal situations and with *structural* properties that are present at the original system.

Then, Rosenblueth and Wiener define closed box and open box models. Closed box models are defined as those in which a fixed finite number of input variables determines a fixed number of output variables, independently of the transition structures. As the model demands a greater number of variables for its full comprehension, it becomes more and more open. Actually, “scientific progress consists in a progressive opening of those boxes” (Rosenblueth and Wiener, 1945, p. 319), leading into more elaborated, sophisticated and hierarchically superior models; both abstract and complex ones.

Finally, Rosenblueth and Wiener argue that imperfection is intrinsically of the nature of models. That is, a one-to-one model would be redundant.⁶

Still on the debate over models, Conant and Ashby (1970) develop a theorem according to which the regulator (of a system) who aims at simplicity and effectiveness should necessarily have a model. “A model *will* be needed” (Conant and Ashby, 1970, p. 90, emphasis in the original). Besides, the regulator must know exactly how the structure of the system works. In other words, the regulator must be able to identify himself with the system

4. About the starting of the scientific process, Rosenblueth and Wiener write: “an intuitive flair for what will turn out to be the important general question gives a basis for selecting some of the significant among the indefinite number of trivial experiments which could be carried out at that stage” (Rosenblueth and Wiener, 1945, p. 317).

5. Rosenblueth and Wiener point out the possibility that material models might be occasionally useless, making them “irrelevant and barren – i.e., a gross analogy” (1945, p. 318).

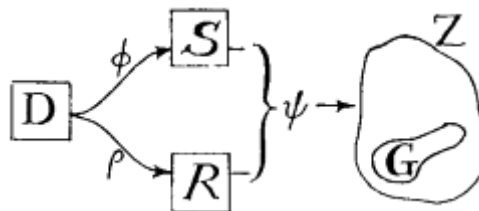
6. See Israel’s (2005) criticisms.

to the point that he knows its structure and its form. Moreover, the continuing effectiveness is dependent upon the deepening of his knowledge of the system.⁷

In order to fully develop the theorem, Conant and Ashby (1970) define: *i*) regulation; *ii*) model; and *iii*) system.

According to the model proposed, regulation (see figure 1) can be defined as: *i*) Z is the set of all events that can occur; *ii*) events are divided into regulated ones (G) or unregulated ones; *iii*) the regulator, in turn, can also generate events (R); what does not hamper the possibility of events happening out of the control of the regulator but within the system (S); *iv*) finally, there are disturbing events (D), or exogenous shocks that interfere in the system; and *v*) if ϕ describes the relationship between D , S , R , the connection DR and ψ , the simultaneous relation among S , R and Z , then R is a good regulator when he achieves the goals contained in G , considering ψ and ϕ , i.e., the system and the shocks.

FIGURE 1
Scheme of regulation



Source: Conant and Ashby (1970, p. 90).

This proposed scheme is not subject to the assumptions of linearity, continuity, temporality or even the existence of a system of measurement. However, when the set of possible events Z is classifiable, the effectiveness of the regulator is higher when his or her own actions are minimal.

The definition of models made by Conant and Ashby cannot be as precise as their definition of regulation because there are times in which modeling implies higher adjustments than that of scale alone. Anyway, it is understood that there is some

7. For a more contemporary approach to modeling, see Page (2012) and Page (2008).

similarity (in the mathematical sense⁸), between ρ and ϕ , i.e., between the influence of shocks upon the action of the regulator in particular and upon the rest of the system. This relationship replaces the mapping of model and system. The proposed theorem is, thus:

the theorem says that the best regulator of a system is one which is a model of that system in the sense that the regulator's actions are merely the system's actions as seen through a mapping h (Conant and Ashby, 1970, p. 96).⁹

The theorem implies that there can be as many effective regulators, whether unnecessarily complex or not. Besides, Conant and Ashby conclude that *the regulator is, in fact, searching for the best model that maps the relations between S and R, conditioned to objective G, which is contained in the universe of system S.*

The suggestions of the authors reinforce the idea that models are compulsory if one aims to understand the mechanisms of the system under analysis. Once the processes are understood, one can move forward and attempt to measure, for example, the efficiency of these processes.

3 PATTERNS, DYNAMIC SYSTEMS, NON LINEARITIES AND CHAOS

May (1976) shows that simple nonlinear models can behave in a vast number of dynamic ways, from stable points to apparently random fluctuations. He considers a model that shows how the magnitude of population (X) in one generation ($t+1$) correlates to the magnitude of that same population in the previous generation (t), which is an example of a first order nonlinear difference equation.

$$X_{t+1} = F(X_t) \tag{1}$$

He highlights that this same logic of recurrence can be found in different fields of science: in biology, for example, it may describe the change of a gene in time; in epidemiology, variable X may represent the fraction of the population infected at each

8. The authors refers to Bourbaki, N. *Théorie des ensembles: fascicule de résultats*. 3rd ed. (Paris: Hermann, 1958) and Riquet, J. Thèse de Paris for the mathematical detailing and they emphasize that the definition of isomorphism can be as diverse as the definition of model. *Bulletin of the Brazilian mathematical society*, v. 76, n. 114, 1948.

9. For the proof of the theorem, please see the original text.

time t ; in Economics it may refer to relations between price and quantity of goods, economic cycles and temporal effects on economic variables; and in the social sciences, it may refer to the bits of information that can be recalled in a period of time t , or to the propagation of rumors in different levels of society.

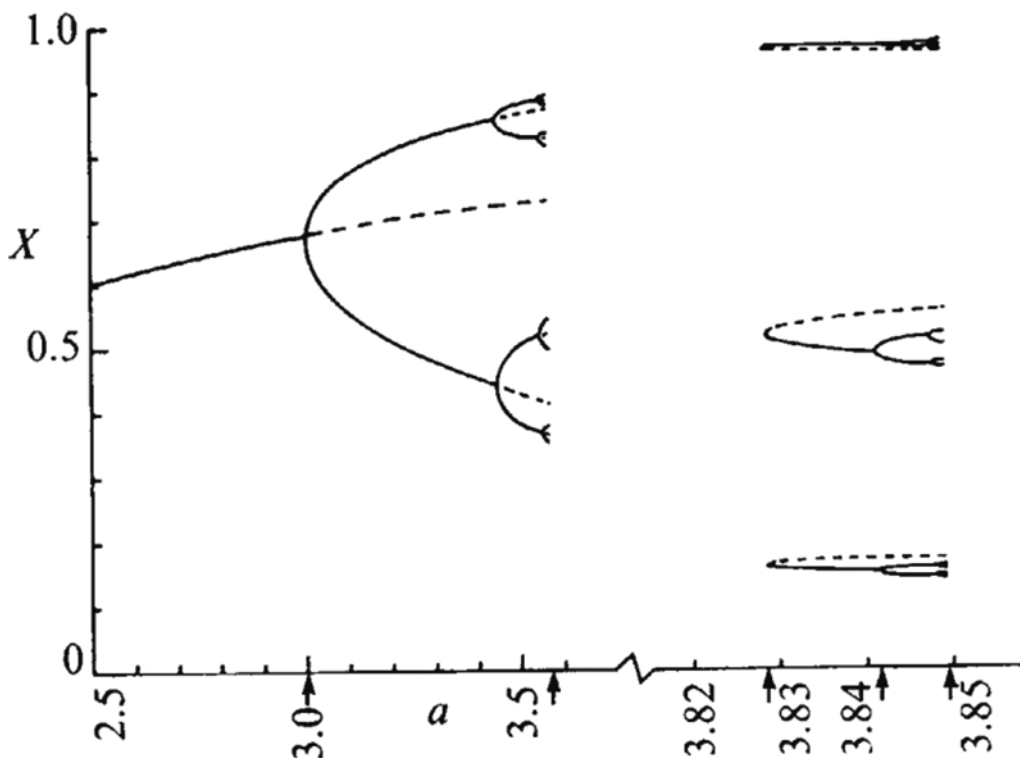
May uses the logistic equation to describe the intertemporal relation between population size in two periods, i.e.:

$$X_{t+1} = aX_t(1 - X_t) \quad (2)$$

Depending on the value of a , the growth parameter, the system can assume different dynamic behaviors, as shown in graph 1.

GRAPH 1

Example of bifurcations using logistic equation



Source: May (1976, p. 462).

The graph illustrates some of the fixed points¹⁰ of convergence of the system for different values of a . In other words, the graph depicts the attractors of the system. The continuous lines are stable fixed points,¹¹ whereas the dotted lines represent unstable fixed points (repellers).

On the left side of the graph, the fixed points are stable up to $a = 3.0$. This means that the population would converge to equilibrium values (fixed points) which would attract all trajectories nearby (stability).

From $a = 3.0$ onwards, the fixed point becomes unstable (note that the line becomes dotted). A pitchfork bifurcation occurs and a cycle of period two starts, followed by a cycle of period four and so on. This means that the population oscillates among these four points, and nearby trajectories are attracted to these cycles. These bifurcations occur a number of times generating what is called 2^n harmonic periods.

When a is above the value of 3.57 none of these cycles are stable. That is, equilibrium values cease to exist and population starts oscillating indeterminably. The point where this occurs is called “point of accumulation of cycles of period 2^n ”, and it establishes the beginning of the chaotic region.

According to Li and Yorke (1975),¹² a chaotic situation occurs when an infinite number of trajectories can occur. Chaos can then be defined as dynamic trajectories that are indistinguishable from a stochastic process.

The right side of the figure shows when a cycle with an odd number of cycles appears, i.e., when population starts to oscillate among three fixed points. This happens by means of a tangent bifurcation and stable harmonics of period 3×2^n appear. Note the accumulation point of stable harmonics of period 3×2^n at $a = 3.8495$. From this point onwards, cycles with all possible integral numbers of periods and also aperiodic trajectories¹³ appear.

10. A fixed point of a function is a number that does not change when the function is applied onto it. That is, x is a fixed point when $f(x) = x$.

11. A fixed point is considered stable when nearby points get closer to the stable point when iterated.

12. Li, T-Y., and Yorke, J. A., *Am. Math. Monthly*, v. 82, p. 985-992, 1975.

13. Aperiodic trajectories never repeat themselves, that is, they do not present periodic behavior.

In sum, what May demonstrates is that a deterministic simple equation of the type of (2) can lead to trajectories that are similar to a random noise. May states that these implications are disturbing.

1. Apparently random fluctuations of a variable (population, for example), occasionally attributed to sampling errors or unpredictable environment may, in fact, come from a rigidly deterministic structure (such as the function of population growth). In other words, processes that are seemingly random may in fact come from a deterministic process. This would be “deterministic chaos”.
2. In a chaotic regime, very small variations in initial conditions may lead to highly diverging trajectories. Thus, even in a system in which parameters are strictly determined, long range predictions may be impossible. This is the classic concept of “sensitivity to initial conditions”, largely described within the theory of dynamic systems.

As a concluding remark May emphasizes that given the ubiquity of nonlinear systems, the study of these systems would be highly important for the adequate training of students:

The elegant body of mathematical theory pertaining to linear systems (Fourier analysis, orthogonal functions, and so on) and its successful application to many fundamentally linear problems in the physical sciences, tends to dominate even moderately advanced University courses in mathematics and theoretical physics. The mathematical intuition so developed ill equips the student to confront the bizarre behavior exhibited by the simplest of discrete nonlinear systems (...). Yet such nonlinear systems are surely the rule, not the exception, outside the physical sciences (May, 1976, p. 467).

Lastly he states that “not only in research, but also in the everyday world of politics and economics, we would all be better off if more people realized that *simple nonlinear systems do not necessarily possess simple dynamical properties*” (May, 1976, p. 467, our emphasis).

The system from which Turing (1952) derives his insights is about chemical substances called *morphogens* or form makers which diffuse and react with the mean where they are. Their diffusion occurs from regions of higher concentration towards those of lesser concentrations according to the coefficient of diffusion of the mean, just like in thermodynamic systems. Turing models the system by simultaneously describing changes by mechanical diffusion as well as diffusion by chemical reactions at the scale of cells and morphogens. Given the rules of transformation and the data, Turing claims that it is possible to identify the individual contribution of each type of change to the change of the system as a whole. The original text of Turing details the parameters and the number of morphogens that render each of the patterns reached (Turing, 1952, p. 50-54).

Turing demonstrates further that given very simple premises, it is possible that homogeneous systems with different rates of diffusion become heterogeneous¹⁴ after receiving small exogenous shocks or by the influence of irregularities in the structure of neighboring cells. Other small changes in the morphogens such as alterations in their chemical concentrations, presence of catalyzers, interfering on the growth of other cells, changes in temperature or in the rates of diffusion may lead the system to become heterogeneous.

Turing also describes how the process of diffusion takes place by stationary or oscillatory, short, very short or long waves. The resulting patterns vary from asynchronous systems to isolated or unstable ones.

Visually attractive applied examples of Turing's proposal, are available at: <http://experiences.math.cnrs.fr/Structures-de-Turing.html>.¹⁵

FIGURE 2

Examples of *morphogens* diffusion of Turing

2A – Initial state



2B – Reached patterns



Source: Chazottes and Monticelli (2013). Available at: <http://goo.gl/sEHrqK>.

Turing summarizes his contribution as just a first analytic step, applied to particular cases, but a step that is dependent only on simple premises such as linearity of the rate

14. The original word used by Turing is *inhomogeneity*.

15. The link was originally put forward by David Feldman in his course *Introduction to dynamics and chaos*, available at: <http://goo.gl/Cr8Vek>.

of diffusion, given by the concentration of the mean. Yet, he explains that it would be possible to include nonlinearities and more concrete biological cases using digital tools.

4 CELLULAR AUTOMATA AND ARTIFICIAL INTELLIGENCE

Wolfram (1984) discusses cellular automata (CA) and the insights that can be drawn from its use. First of all, Wolfram defines CA as simple discrete dynamic systems that *have properties of self-organization*. Fundamentally, CAs differ from traditional dynamic systems by the emphasis on spatially rigid rules, that is, the local influence of a given neighborhood. However, just as in the analysis of dynamic systems, the interest falls onto the trajectories of the CA system and in its attractors.

Formal characterization of a CA is made by two parameters: r which determines the distance of influence of the neighbors from each cell into the next state; and k , which represents the number of states that each cell can have. So, the basic form of a CA corresponds to a $k = 2$ and $r = 1$, configuring the neighborhood as the immediate (one cell distance) neighbors of cell i . Values of k and r determine the number of possible rules to be applied to the system, according to equation: $Rules = k^{k^{(2r+1)}}$, so that the simple case contains 2^8 or 256 rules.¹⁶

Given the definition and processes of CA, Wolfram identifies four typical classes of behavior. Three are similar to fixed points, cycles and chaos (i.e. they have strange attractors) and are also found in the analysis of dynamic systems. The fourth class may be capable of universal computation capacity. Qualitatively speaking, Wolfram identifies the following classes:¹⁷

- homogeneous states: from a finite number of evolutions (from one state to the next), for all initial states, cells assume a single state;
- periodic: typically with small periods;
- chaotic patterns; and
- complex structures.

16. The description of the code used to nominate each rule is proposed by Wolfram in 1983 and it involves the sorting of possible rules and its transformation into decimal notation.

17. Wolfram also distinguishes the classes quantitatively using a spatial entropy measure and the verification of the initial random state and its deviation from randomness with different probabilities through time.

The first class allows predictions to be made with probability one, independently of initial conditions. In class 2, the result of a specific cell can be given as long as its initial state is known. Class 3 leads to random states. However, from the initial states, it is possible to build an algorithm that leads to the correct prediction. Class 4 does not allow for predictions. The algorithm necessary to make such prediction would be as complex as the same simulation of the evolution of the CA's states. In this case, a model is not possible.

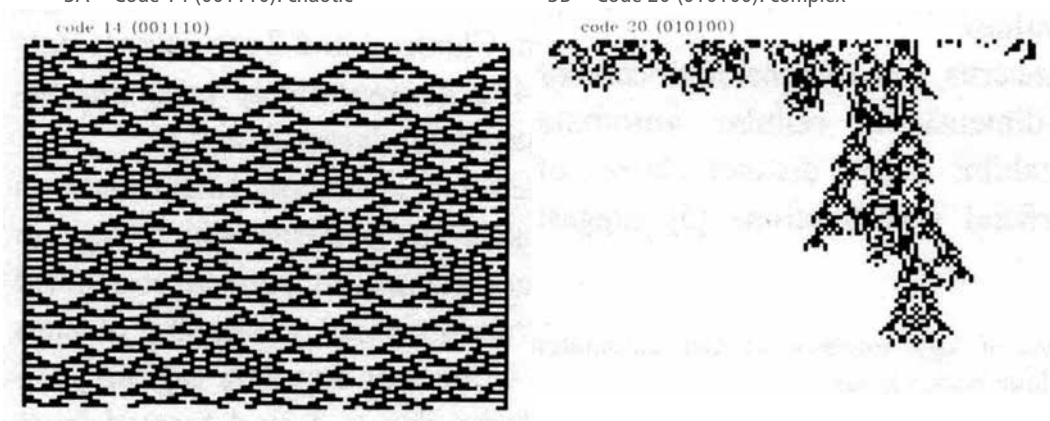
Further qualitative difference from chaotic and complex classes can be seen in the comparisons of rules 14 (001110), chaotic (figure 3A) and 20 (010100), complex (figure 3B).¹⁸

FIGURE 3

Examples of cellular automata with chaotic and complex patterns

3A – Code 14 (001110): chaotic

3B – Code 20 (010100): complex



Source: Wolfram (1984, p. 6).

Wolfram reaffirms that the universality property may very well be present in cellular automata:

Cellular automata may be viewed as computers, in which data represented by initial configurations is processed by time evolution. Computational universality implies that suitable initial configurations can specify arbitrary algorithms procedures. The system can thus serve as a general purpose computer, capable of evaluating a (computable) function. Given a suitable encoding,

18. In a later text, Wolfram (2002) details the difference between classes 3 and 4. While class 3, the complex one, presents “complicated” and random behavior; class 4, the chaotic, can be called a mixture of chaos and order, processes that emerge and collapse, processes that happen locally and not necessarily are transmitted to the rest of the system.

the system may therefore in principle simulate any other system, and in this sense may be considered capable of arbitrarily complicated behavior (Wolfram, 1984, p. 31).

Wolfram concludes that “cellular automata are simple in construction, but are capable of very complex behavior” (Wolfram, 1984, p. 34). Moreover, he confirms that the four classes discussed are present in all unidimensional cellular automata.

The interest of studying CA, from the view of physicists, is that CAs are characterized as discrete dynamic systems in space-time and were the starting point for all the following analysis about phase space (Langton, 1986). In fact, this is how Langton comments von Neumann’s works: “Von Neumann proved that cellular automata are capable of universal computation by showing that a universal Turing machine could be embedded in a cellular array” (Langton, 1986, p. 129).

4.1 An example in biology

Langton (1986) makes a bold proposal in his paper discussing the “possibility that life could emerge from the interaction of inanimate artificial molecules” (Langton, 1986, p. 120). He does that using the properties of CA and its ability to reproduce “molecular logics”; a term originally used by Lehninger.¹⁹ Langton proposes to describe the functional role of the molecules first and then the possibility that a virtual system of automata can reproduce that functional role. Thus, Langton makes two systems that should reflect typical behavior of living organisms.

Given the four larger groups of biomolecules (proteins, nucleic acids, polysaccharides and lipids), the first two play the most relevant role. In fact, such is the order of magnitude of this relevance that, in practice, proteins effectively control whether a given reaction occurs or not. These roles are: *i*) catalysts; *ii*) transport; *iii*) structure; *iv*) regulatory (and hormonal); *v*) defense and immunodeficiency; and *vi*) information database, memory.

According to Langton, the necessary properties for computer programs to be able to reproduce the observed dynamics at the molecular level are logically simple. In order to do so, it is enough that an “initial set of operators and operands goes to work producing more

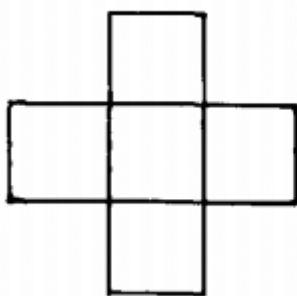
19. Lehninger, A. L. *Principles of biochemistry*. Worth: New Your, 1958.

operators and operands, which immediately enter into the ongoing logical fray” (Langton, 1986, p. 122). Computationally, it is enough that the computer withholds interactions many-to-many and that it allows the change of an operator by another one. That is, that the modification of a program and its data can be made by another program. CAs, as described by von Neumann (1966), have these properties. Langton suggests the use of CA in two dimensions with a transition function Θ and two kinds of neighborhoods.

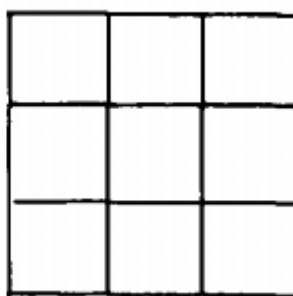
FIGURE 4

Original example of types of neighborhoods for cells

4A – Five cell neighborhood



4B – Nine cell neighborhood



Source: Langton (1986, p. 123).

Langton proposes two ways to study CAs. The first, alongside with von Neumann, would be to specify a rule Θ and verify its behavior; the other would be to specify a behavior and identify the Θ that generates that behavior. Besides, Langton refers back to the classification of evolution made by Wolfram and emphasizes the possibility of connecting results observed in CA in relation to those on the physic and biological worlds. Especially, Langton focuses on the fact that both systems often reach cyclic patterns with clearly defined strange attractors.

Langton also highlights the “emerging properties”, i.e., that the property not explicit in rule Θ emerges from local influence.²⁰ It is from this original idea that the notion of “bottom-up approach” is derived. In order to defend the argument of

20. Langton explains: “the global behavior of a society is an emergent phenomenon, arising out of all of the local interactions of its members (...) from our previous discussion (...) we know that complex behavior can emerge from the interaction of very simple living parts” (Langton, 1986, p. 134). This also implies the notion of the butterfly effect or “sensitivity dependence on initial condition”, quoting Wilson, “a slight alteration [of the termite’s response] to a particular structure will tend to be amplified to a much greater alteration in the final product” (Langton, 1986, p. 135, *apud* Wilson, 1975, p. 12).

emergence, Langton defines parameter λ as the ratio of the “number of neighborhood states that map to a non-quiescent state/total number of neighborhood states” (Langton, 1986, p. 126). The variation of parameter λ enlightens the analysis of the results. The author proposes an analogy with temperature scales. For small values of λ , the dynamics of CA are stable with little changes; for high values the dynamics are unstable such as a heated gas; for medium values of λ (in the region the author calls region 2), the interesting dynamics show up. In fact, it is by using the periodic structures and balanced CAs that Langton searches for the possibilities of artificial life.

von Neumann set out to demonstrate the existence of a Turing machine that could effect its own reproduction. Since he was able to demonstrate that such a machine can exist, it becomes plausible that many, perhaps all, of the processes upon which life is based are algorithmically describable and that, therefore, life itself is achievable by machines (Langton, 1986, p. 136).

The concept of Turing machine is used by Langton to propose a Virtual Turing Machine (VSM). The Turing machine – which according to Langton is comparable to a CA and is a predecessor of computers – would be a theoretical program that can read and write (save and retrieve) information in an endless extendable tape containing a group of rules that states which information is to be saved in the following moment in time ($t + 1$). The Virtual Turing Machine (VSM) proposed by Langton would involve CAs of order two or higher, that is, those that are not present in the initial transition rule Θ and, that, moreover, are transmitted to more than one cell. According to Langton, VSMs would have the properties to be at the same time data and process, as they are written on the own tape where they operate. They could be able to write (build), erase, self-erase, reproduce (make other VSMs), read and modify structures, besides being able to use other VSMs memories. It is from this list of VSMs’ properties that Langton derives the functions of biomolecules. Going further, Langton emphasizes the relevance of the property of VSMs to be able to replicate themselves (just as DNA). He suggests that it would be possible to make a colony of machines that would interact and evolve as time passes by.

When explaining that his intention is to replicate the roles of molecules and not the molecules themselves, Langton coins the term artificial life, which would later gain prominence.

Langton makes the connection back to biology and uses VSMs' properties to illustrate the case of ants and artificial biochemistry. He concludes that there is some "arbitrariness" between the complex phenomenon that is observed and the simplicity of the rules of the hardware used.

John von Neumann (1966) starts his essay on *The role of high and of extremely high complication* pondering the difference between artificial automata – computers – and those said to be "natural" – living organisms, especially, humans.

He says that in terms of size, human brain would have about 10^{10} neurons. After some calculations, he reflects that in general "the nervous system outperforms the machine by a factor of roughly 200" (1966, p. 65). Von Neumann evaluates, however, that the possibilities of achievements is really in the interconnections of the components of the system. Following along in his comparison, he states that from the point of view of physical organisms, it is not possible to be sure about the implicit mechanisms of how memory is established.

Another relevant difference von Neumann points out is that living organisms constitute a system that is so integrated that it can run in spite of the occurrence of errors. Von Neumann describes human beings as similar to a complex system of self-organization:²¹

The system is sufficiently flexible and well organized that as soon as an error shows up in any part of it, the system automatically senses whether this error matters or not. If it doesn't matter, the system continues to operate without paying any attention to it. If the error seems to the system to be important, the system just blocks that region out, by-passes it, and proceeds along other channels. The system then analyzes the region separately at leisure and corrects what goes on there, and if correction is impossible the system just blocks the region off and by-passes it forever (von Neumann, 1966, p. 71).

Von Neumann concludes that it is not possible to compare natural and artificial automata if one does not know how the "complicated mechanisms" of the components' interact. Further, it is necessary to consider the performance of each system given each

21. He cites the capacity of organisms to self-observation and self-organization, probably associated to its autonomy.

kind of problem. The context is central to determine the efficiency of the analysis and the statistics by themselves are not enough to provide adequate answers.

The original text by von Neumann is divided in lectures. However the text that we were able to access contains only lectures 4 and 5. The lectures are commented by Arthur W. Burks who contextualizes some of the propositions of von Neumann. The fifth lecture is entitled *Re-evaluation of the problems of complicated automata – problems of hierarchy and evolution*.

In fact, as criticized by Israel (2005),²² von Neumann does not distinguish between “complication” and complexity. The author defines complexity saying that “an object is of the highest degree of complexity if it can do very difficult and involved things” (von Neumann, 1966, p. 78).²³

Von Neumann reinforces his description of living organisms as “complicated aggregations”, probabilistically “improbable” amalgamation of elementary parts. Implicitly, we may say that he is agreeing with Anderson (1972). Von Neumann goes further and describes his own surprise with the fact that organisms can generate other organisms that are even more “complicated” than themselves, and that the original recipes are not necessarily within the original recipes, with no hints or “predictions” about how the succeeding organism should be. This does not occur when artificial automata is at play. That is, the synthesis made by the automata must be entirely described.

Going back to the definition of “complication”, von Neumann suggests that one way to quantify complication is by the number of constituent parts. He then establishes the concept of “critical size” which would be the point after which still not observed properties (such as the capacity to generate successors that are more complex than the antecessors) would emerge.²⁴ By the end of the text, in a proposal to actually build an automaton, von Neumann proposes the notion of “mutation” to a computational element. This mutation could be lethal for an automaton, but could also be inherited.

22. See section 8.

23. Note that the meaning of involved in this case is of difficult to comprehend, complicated, intricate, complex, elaborated and not involvement.

24. Our choice of emerge in this context.

Turing (1950) uses the word machines to refer to electronic or digital computers that follow a book of rules from which they cannot deviate. He describes digital computers as those that contain three parts: storage, executive unit and control. Storage refers to the place where information is stored and is equivalent to the paper where humans take notes. The executive unit does operations that are involved in calculus. Control guarantees that the machine follows the rules that had been previously established.

Moreover, digital computers would be machines in discrete as opposed to continuous state. Given an initial state and given the inputs into the system, the machine can predict all future states.²⁵ Thus, a digital computer would be able to mimic the behavior of any other digital machine. This important property is called universality. The machines engineered by Turing are thus called universal Turing machines.

A consequence of this universality is that all digital computers are in a way equivalent. So, the question whether machines think could be reduced to “is there any machine of discrete state that it would do well in an imitation game?”.

Turing (1950) speculates that such a machine would be feasible in a 50-year horizon. About the philosophical question of whether machines do think (and thus learn), Turing introduces some points for discussion that would eventually lead to artificial intelligence.

Marvin Minsky (1961) reviews the state of the art of the field of artificial intelligence at the time. He focuses on the classes of activities that a generic computer is programmed to do and that leads to a superior level of processing, such as learning and problem solving.

Minsky discusses the first ongoing steps toward artificial intelligence. He divides the problem of heuristic programming in five main areas: search, pattern recognition, learning, planning and induction, and defines the concept of heuristics as an adjective,

25. Turing emphasizes that the property of being able to predict future states from knowledge of initial state restricts machines to a discrete state. He hints at the *butterfly effect* stating that “the system of the “universe as a whole” is such that quite small errors in the initial conditions can have an overwhelming effect at a later time. *The displacement of a single electron by a billionth of a centimeter at one moment might make the difference between a man being killed by an avalanche a year later, or escaping*” (Turing, 1950, p. 440, our emphasis).

related to the improvement of performance in problem solving and, as a noun referring to any method or trick to improve the efficiency of a system in problem solving.

Minsky says that a heuristic program should be successful in a variety of problems, but not necessarily in all of them. The program is considered successful when there is a general improvement in the performance of problem solving.

Minsky examines problems that are self-defined, i.e., for which there is a systemic way to define when a proposed solution is acceptable. Besides, all problems considered are trivial, i.e., it is possible to find a solution – if one does exist – by the exhaustive search among all possible solutions.

Thus, goes Minsky, the first stage for building artificial intelligence consists in search. In order to solve a problem, a machine can be programmed to test a vast space of possible solutions. For trivial problems, this process is very effective, but highly inefficient. The heuristics of search would consist in finding techniques that would allow using incomplete results of the analysis to make the search more efficient.

When there is some knowledge of the search space, hill-climbing algorithms may be useful. In this case, the process of search becomes a simple case of local optimization. Using such an algorithm, one starts from a random point, examines the surroundings and moves toward the highest inclination point. The same process is repeated until it becomes impossible to increase the value of the variable one wants to maximize. Some problems of this approach consist in reaching only local optimum, when the goal is to reach a global maximum; or yet, to reach a flat table, the so-called mesa phenomenon.

Pattern recognition constitutes another heuristic resource, by which the machine would classify problems according to the most efficient method of solution. In other words, the machine would be able to identify patterns of problems that would be more efficiently solved by specific method.

In order to do that, the machine would have to identify the relevant heuristic characteristics of the object at hand. Some of the methods for that consist in linking objects to models or prototypes; and testing each one to identify such relevant heuristic characteristics. In complex problems, however, Minsky (1961) notes that it may be

challenging to divide complex objects into parts and describe the complex relationships among them.

The machine would also be able to increase its efficiency in problem solving if it could consider its own previous experience. The learning heuristics imply that the machine would apply similar methods that had previously worked for similar problems. The implementation of these systems is based on decision models that positively award past successes. However, in complex systems it may be hard to tell which decisions were relevant to find the solution.

Minsky also mentions the cost of unlearning. The machine will built-up its memories based on past experiences and, in case there are better procedures than the ones experienced, the machine would incur the cost of correcting its old “beliefs”.

The problem of extinction of “unlearning” is especially critical for complex, hierarchical, learning. For, once a generalization about the past has been made, one is likely to build upon it. Thus, one may come to select certain properties as important and begin to use them in the characterization of experience, perhaps storing one’s memories in terms of them. If later, it is discovered that some other properties would serve better, then one must face the problem of translating, or abandoning, the records based on the older system. This may be a very high price to pay. One does not easily give up an old way of looking at things, if the better one demands much effort and experience to be useful. Thus the training sequences on which our machines will spend their infancies, so to speak, must be chosen very shrewdly to insure that early abstractions will provide a good foundation for later difficult problems (Minsky, 1961, p. 30-31).

In order to solve complex problems, the machine has to manage a set of interrelated sub problems. A step by step heuristic would imply in: subdividing the problem; selecting sub problems to be solved (estimating relative difficulty; estimating the centrality of different sub problems); and selecting adequate methods for the solution of the sub problems.

However, for difficult problems, the machine would have to analyze the problem as a whole, that is to plan. A number of planning schemes are presented by Minsky, which include the use of analogy, semantic or abstract models.

Minsky compares the study of neural networks with the field of artificial intelligence. While the former aims at demonstrating that heuristic simple processes can be made

of a collection of simple elements with no organized relationships, the former looks for more powerful heuristics to solve very complex problems.

The work on “nets” is concerned with how far one can get with a small initial endowment; the work on “artificial intelligence” is concerned with using all we know to build the most powerful systems that we can” (Minsky, 1961, p. 53).

Enabling machines to solve complex problems means that they should have inductive capacity, that is, they should have methods that make it possible for them to build general statements about events beyond their own experience. But for machines to answer questions about hypothetic events beyond their own experience, without trying out those experiments would mean that the answers would have to come from an internal submachine which is within the original machine. The problem of inductive inference can thus be seen as the reconstruction of this internal machine; and this internal machine can be understood as a model that regulates the world. Since the machine is part of the world, the internal model would have to include representations of the machine itself.

If one asks the creature “why did you decide to do such and such” (or if it asks this of itself), any answer must come from the internal model. Thus the evidence of introspection itself is liable to be based ultimately on the processes used in constructing one’s image of [the creature on] one’s self (Minsky, 1961, p. 57).

5 INFORMATION THEORY

The information theory discussed in this section brings original elements of computing and even of the CAs discussed in the previous section. Information in this context superimposes the notions of entropy and uncertainty. In fact, information may be considered as a measure of complexity of a system. We initiate the section with a brief concept of entropy from thermodynamics, followed by the discussion over the paradox of Maxwell’s devil. We then use Bennett’s paper (1973) to approximate the notions of entropy and information. Bennett also puts forward the concept of reversible computation. Herbert Simon (1973) then discusses the definition of nearly decomposable, introduces the alphabet as a constituent element of language and the notion of loose coupling. After all these concepts, we present information theory as proposed by Shannon (1948).

5.1 Maxwell's devil and entropy

According to the second law of thermodynamics, in an isolated system, entropy always increases, up to its maximum value which consists in absolute disorder. Entropy increases because when energy is transformed in work, there is loss of heat and this loss characterizes entropy.

Maxwell's devil paradox tried to show that the second law of thermodynamics possessed only statistical certainty, being possible, hypothetically, to violate such law. Maxwell supposes an isolated container where there are fast and slow particles in movement. A devil controls the door that divides the container in two. Such devil would open the door so that fast particles coming from the right could enter the left part of the container. After some time, the system would be more organized than it was in the beginning with a concentration of faster particles on the right side and of slower particles on the left, thus reducing entropy and contradicting the second law. According to Maxwell, such phenomenon would happen only by the wit of an attentive being – the devil.²⁶

Bennett (1973) showed, however, that it is possible to observe and retain information without increasing entropy. In other words, he showed that any computation could be done without the use of energy.

In order to do that, Bennett introduced the mechanisms of *reversible computing*. In the first stage, a logically reversible automaton would save all intermediate results of the process of computation so as to avoid the irreversible erasing of data. In the second stage, the automaton prints the obtained results. In the third stage, it gets rid of all intermediate results, redoing all steps of computation in reversible order and thus taking the machine back to its original state. In the final stage, the machine would have the desired final result without maintaining unwanted information which would demand work to be erased.

26. "The hot system has got hotter and the cold colder and yet no work has been done, only the intelligence of a very observant and neat-fingered being has been employed", citation of Maxwell from Mitchell (2011, p. 45).

Landauer (1961) showed that any irreversible logic operation of information manipulation, such as erasing a bit of memory, causes an increase of entropy and dissipates an associated amount of energy as heat. According to the line of thought developed by Bennett, in order to separate the fast particles from the slow ones, at some moment, Maxwell's devil would have to delete its memory as the lack of space to store information would ensue. The paradox would thus be invalid giving that the act of deleting memory increases entropy by the same amount that it would have been reduced by the organization of the particles.

It can be said that Maxwell's devil continues to be controversial (Mitchell, 2011). Anyhow, these debates provided the foundation for the development of information theory, detailed below.

5.2 Nearly decomposability

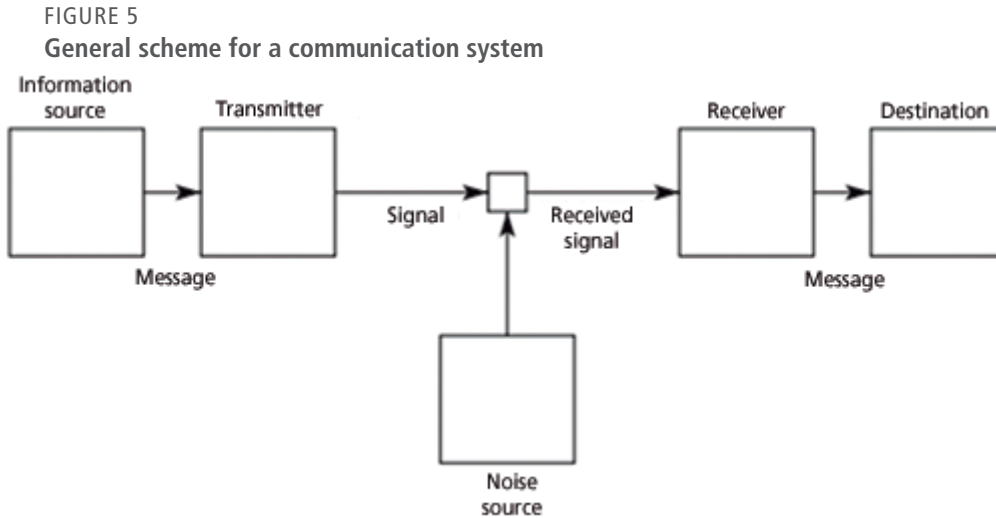
Simon (1973) introduces the concept of nearly decomposability, according to which it is possible to distinguish the interactions of a system using the frequency of interaction among its components. If the interactions observed among the particles of a system occur less frequently than a given time T , then those interactions can be considered constant. Similarly if the interactions occur with a higher frequency than another given parameter τ , the interactions can be disregarded as nuisance.

Another innovative contribution of the work of Simon is the notion that the alphabet is the key analogic element for the construction of complex systems: letters composing words that, in turn, in a hierarchical sense, compose language *per se*. From this notion, Simon poses that given the limited number of options, the evolutionary process is more effective. For example, milk's protein is composed of only five of the 92 natural elements available. That is, a few interactions are enough, as long as they are meaningful.

Simon also contributes with the insight of *loose coupling*, according to which the independence among subsystems of a higher system implies functional effectiveness, in the sense that each subsystem is autonomous and self-organizes in the absence of other subsystems.

5.3 Information theory

Shannon's paper (1948) establishes definitions that have become classic in communication theory.



Source: Shannon (1948, p. 624).

Obs.: information source: produces the message or messages to be transmitted in various formats, such as numbers, audio, positions; transmitter: converts the original message so that it can be transmitted over the chosen channel, such as transforming voice into electrical current; channel: mean used so that the message can travel from emitter to receiver; receiver: converts the message back to its original format, for example, electrical current back into voice; and destination.

It is also clear by the analysis of figure 5 that the path of the message through the channel of transmission can suffer unintended interference of noise, as detailed below.

Shannon defines channel capacity as:

$$C = \lim_{T \rightarrow \infty} \frac{\log N(T)}{T} \quad (3)$$

where $N(T)$ is the number of signs sent of time length T and C is the maximum rate of transmission of information for a given channel, even in the presence of transmission error generated by noise.

Shannon then describes the source of information as a symbol generator in a stochastic process that can be successively detailed in order to be approximated by natural language. In an example, he suggests to use words of English associating them to probabilities of occurrence. This stochastic process is described in the literature as

discrete processes of Markov and can be defined as such: for a given finite set of states S_p , associated to a set of probabilities $p_i(j)$, the system passes from one state S_i into another state S_j with probability $p_i(j)$. Note that this definition is similar to that of CA proposed by Wolfram (1984).

Shannon's next step is to propose a measurement of the rate of production of information by a given source. His definition of rate of production of information coincides with that of Kolmogorov (see below) and also with Boltzmann's statistical mechanics.²⁷ Starting from the definition of entropy, Shannon (1948) defines also "relative entropy" as the entropy divided by the maximum possible entropy. Thus, he defines *redundancy* as the unit less the relative entropy.

In this format (see figure 5), the efficiency of the system when coding a message to be sent through a channel equals the ratio between the rate of codifying and the actual capacity of the channel. Shannon then defines distortion as what occurs when the signal received is a defined function of the signal sent, that is, when the difference between the signal that was sent and that was received is systematically the same. Distortion has the advantage that it can be fixed.

Noise occurs when the difference between signal sent and received are not always the same. In these cases, it is not possible to rebuild the signal sent with precision. In a statistical sense, a system with noise contains two statistical processes in action: the source one and the noise one (see figure 5).²⁸ Redundancy as defined above allows the probability of errors to be small.

27. Boltzmann defines entropy of a macrostate as a function of the number of microstates that could have led to those macrostates. If there is no work, Boltzmann's entropy increases until it reaches the macrostate with the higher possible entropy. See a simple example. A gambling machine contains the combinations of possible results, three by three, of bananas, lemons, apples or cherries, the microstates. The macrostates refer to "win" or "lose". The numbers of microstates that are associated to lose are much larger than the number of microstates associated to win. Thus, Boltzmann's entropy of macrostate lose is much higher than the entropy of lose. The representation of Boltzmann's concept is written in his tomb: $S = k \log W$, where S is entropy, W is the number of microstates possible corresponding to a given macrostate and k is Boltzmann's constant, used to represent entropy in standardized units (Mitchell, 2009, p. 50-51).

28. The author defines "equivocation" as the conditioned entropy which measures the "average ambiguity of the received signal" (Shannon, 1948, p. 399).

Kolmogorov (1965) proposes the definition of complexity as being the measurement of computational resources necessary to describe an object, in other words, the smallest program capable of specifying a given object. In order to illustrate this concept, imagine “64 zeros”. Only eight characters were necessary to describe such a piece of information, which in its full form would have taken exactly 64 characters. In contrast, an information apparently aperiodic, such as “000111000111010000100011111010” cannot be described with a program that has at least the same length of the original object.

Kolmogorov (1965) proposes further a new approach to describe information quantitatively beyond the paradigms of combinatorial and probabilistic analysis: the algorithm approach. According to him, combinatorial approach calculates the possibilities of the construction of “words”, given the “letters” of the alphabet. Probabilistic analysis is possible, but it can lead to senseless results – such as the possibility of negative results for entropy – or results difficult to comprehend. Take the case of literary work. Using *War and Peace* by Tolstoy as an example, Kolmogorov states: “Is it reasonable to include this novel in the set of ‘possible novels’, or even to postulate some probability distribution for this set?” (Kolmogorov, 1965, p. 3).

Then, Kolmogorov defines relative complexity using his proposed algorithmic approach to be “the minimal length $l(p)$ of the ‘program’ p for obtaining y from x ” (Kolmogorov, 1965, p. 5).

$$\varphi(p, x) = y \tag{4}$$

Thus, relative complexity of $(K)^{29}$ is such that:

$$K_{\varphi}(y/x) = \begin{cases} \min_{\varphi(p,x)=y} l(p) \\ \infty, \text{ if there is not } p, \text{ such that } \varphi(p, x) = y \end{cases} \tag{5}$$

Within the context of information theory, *entropy is the simultaneous measure of the uncertainty of a given variable and the information content that is present*. So that the more content there is in the message, the greater the uncertainty. Specifically,

29. The author shows that the values for relative complexity are always positive and highlights the fact that text is not related to an assumed complexity in creating the program p . This demonstration is also present at Shannon (1948).

Shannon's entropy – also known as statistical entropy – is a measure of the quantity of information present in a message, or, in other words, the mathematical expectation: the chance that a given expected result occurs. So Shannon's entropy is maximum when all results are equally probable and it decreases when other results have higher probability. As an example, it can be said that entropy is 0 when there is one possible result with probability one. In the case of a non-biased die of six faces, entropy is 1,79.³⁰

Gell-Mann and Lloyd (2004) offer a definition of effective complexity of an entity as the “length of a highly compressed description of its regularities” (Gell-Mann and Lloyd, 2004, p. 387), in other words, the smallest complete description of the patterns of a given object or system. Detailing the definition further, the authors refer to the concept of information theory. In a summarized way, information Σ is the sum of the complete set of Y plus the probability distribution of the predictions of the theory among the possible sets of data I , so that:

$$\Sigma = Y + I \quad (6)$$

Formally, Gell-Mann and Lloyd call Y the *Algorithmic Information Content* (AIC) of the set, and I , the contingent average of the AIC. Then, they define the “AIC of a bit string (and, hence, of the entity it describes) [as] the length of the shortest program that will cause a given universal computer U to print out the string and then halt” (Gell-Mann and Lloyd, 2004, p. 388). Note that Kolmogorov (1965) and Chaitin (1966) have also proposed AIC as a measure of complexity. However, it is relevant to say that strictly speaking AIC is not computational itself, but it is only used as a theoretical tool to support the thinking.

Still according to Gell-Mann and Lloyd, the intention of the scientist when proposing a theory is to minimize both terms on the right of the equation so that the theory remains simple (low Y) and at the same time accurate in relation to the data. Thus, the *tradeoff of the scientist is to include as much detail as possible in order to make more precise predictions, at the cost of complexifying its theory*.³¹

30. A die of 20 faces, at a higher level of uncertainty, has entropy of approximately 2.996.

31. Another way to interpret $Y + I$ is to identify the first term as the basic program, the rules, whereas the second term represents the data needed.

Gell-Mann and Lloyd (2004) also define pseudo complexity for the cases when there is not a generator program, such as a CA with complex results. That is necessary because the reduction is not the AIC, as previously defined.

6 NETWORKS

Hopfield (1982) shows that computational properties such as the stability of memories and the construction of categories of generalization can emerge as an spontaneous result of the interaction of a large number of simple neurons. In other words, he shows that computational properties emerge as collective properties of systems with large number of simple equivalent components, such as neurons. “Any physical system whose dynamics in phase space is dominated by a substantial number of locally stable states to which it is attracted can therefore be regarded as a general content-addressable memory” (Hopfield, 1982, p. 2.554). In those systems it is possible to access information that is stored in the memory via their attributes (a simple example: the information “banana” could be accessed via attributes such as fruit, yellow, and monkey).

Thus, Hopfield proposes a neural network in which the state of each neuron changes randomly in time and asynchronously:

- each neuron i can be in each of two states $V_i = 0$ (not firing) and $V_i = 1$ (firing at maximum rate);
- the strength of synaptic connection between two neurons i and j is given by T_{ij} ;
- a neuron is not connected to itself;
- the state of the system is given by the N values of V_i , where N is the number of neurons;
- the state of the system changes according to the algorithm below. Note that each neuron processes information in parallel, but not simultaneously:
 - a) each neuron has a fixed limit U_i ;
 - b) each neuron readjusts its state randomly in time, but with an attempt rate W , that follows the rule

$$V_i \rightarrow 1 \text{ if } \sum_{j \neq i} T_{ij} V_j > U_i \quad (7)$$

$$V_i \rightarrow 0 \text{ if } \sum_{j \neq i} T_{ij} V_j < U_i \quad (8)$$

c) the stochastic mean processing time is $1/W$.

This model uses strong non-linearities to make choices, categorize and regenerate information. Hopfield states that the analysis of collective emerging effects and of spontaneous computation should necessarily emphasize non-linearities and the relationship of inputs and outputs.

Hopfield considers the special case in which $T_{ij} = T_{ji}$ (symmetry) and defines Energy as a monotonically decreasing function:

$$E = -\frac{1}{2} \sum_{i \neq j} T_{ij} V_i V_j \quad (9)$$

Thus, changes in the state of the network occur until a local minimum is reached.

Simulations of the model show that the algorithm has the necessary properties to constitute a content-addressable physical memory. Memories are stored as stable entities that can be recalled from a large enough fragment and occasional ambiguities are solved statistically.

Other emerging collective properties detected in the simulation were: capacity of generalization, categorization, error correction; and time sequence retention. Hopfield also states that the collective properties do not depend on the accuracy of the details of the modeling or on the failure of individual devices.

Finally, Hopfield suggests that the use of a similar model using integrated circuits would lead to chips more reliable and less sensitive to element failure and soft-failure than regular circuits; and that the capacity of asynchronous parallel processing of the model would lead to quick solution of specific computational problems.

In sum, Hopfield develops a model of neural networks in which a large number of neurons perform parallel and asynchronous processes. From the collective behavior of these simple processors, new computational properties emerge and the system constitutes itself as a content-addressable memory. Those memories can be accessible given large enough fragments. Given its structure as a combination of parallel asynchronous processing elements, the system is robust to failure in individual devices and behaves in tandem with the concept of loose-coupling as defined by Simon (1973).

Bernard Derrida, Gardner and Zippelius (1987) relax the hypothesis specified by Hopfield that the synapses are symmetric and that each neuron connects itself to an infinite number of other neurons. They prove that such a model can be solved exactly, even under those conditions. The hypothesis assumed by Derrida, Gardner and Zippelius are more realistic as the synaptic connections observed in biological networks are asymmetric and that each neuron in fact connects itself to a small fraction of all neurons.

Derrida, Gardner and Zippelius also obtain the analytical expression for the evolution of a configuration with a finite superposition in a stored pattern, i.e., of a configuration that consists of the fragment of a stored memory. Even when the systems recalls the information, two close enough initial configurations do not configure the exact same memory. This fact suggests that the attractor that is near a stored pattern possesses such a structure that is more complex than a unique attracting configuration. In other words, when the system is given familiar information, it ends up recalling the original information. However, the memory retrieved as a consequence of different familiar information is not exactly the same. Further, when two stored patterns are correlated (correlated memories), there are regimes by which the system can recall these patterns but it cannot distinguish one pattern from the other.³² Derrida, Gardner and Zippelius suggest the generalization of their approach to stored patterns that are time-dependent.

7 EVOLUTION

How can we explain the increase in the complexity of living organisms, given the second law of thermodynamics?³³ In other words, given that entropy in the universe always increases, how have living systems become more organized and complex?

Darwin's theory of evolution provided the most compelling explanation. According to the theory:

The result of evolution by natural selection is the appearance of “design” but with no designer. The appearance of design comes from chance, natural selection, and long periods of time.

32. Suppose two siblings with similar appearance (correlated stored patterns) are introduced to a new person. Later, when this person meets either of the siblings it is possible that the person correctly recalls which of the sibling he or she is meeting. However, it is also possible that the person cannot distinguish which of the siblings it is.

33. See section 5.

Entropy decreases (living systems become more organized, seemingly more designed) as a result of the work done by natural selection. The energy for this work comes from the ability of individual organisms to metabolize energy from their environments (e.g. sunlight and food) (Mitchell, 2011, p. 79).

Thus, biological evolution is the greatest proof that simple rules – such as the algorithm of natural selection – are able to generate highly complex products with no need of a designer (or a central control).

This Darwinian vision was combined with Mendel's vision to lead to what is today called the *modern synthesis* of the theory of evolution. The modern synthesis became largely accepted by biologists and can be characterized by the following principles concerning evolution:³⁴

- natural selection is the main mechanism of adaptation;
- evolution is a gradual process that occurs via natural selection of small variations in individuals. These variations occur randomly by mutation and genetic recombination and do not follow a biased trend, i.e., not all of them lead to improvement; and
- macro scale phenomena, such as the origin of new species, can be explained by microscopic processes of genetic variation and natural selection.

In the 1960s and 1970s, the validity of the modern synthesis was challenged. A study with fossils by Gould and Eldredge (1977) verified some discrepancies between the theory proposed by the modern synthesis and the evidence found. These paleontologists suggested that punctuational change – not gradual ones – dominated the history of life on earth. According to them, evolution would be concentrated in rapid events of speciation³⁵ and would not occur in a constant rhythm in tune with geological scale. Their observations of fossils showed that most organisms remain practically unaltered through geological time and that evolutionary changes tend to be concentrated in rapid events of speciation. This somewhat antagonistic view to the gradualism established by the modern synthesis became known as theory of punctuated equilibria.

34. Based on Mitchell (2011).

35. Evolutionary process from which new biological species appear.

Gould and Eldredge moved beyond the field of biology suggesting that their theory of punctuated equilibria could constitute a general philosophy of change. As an illustration of the adherence of such a theory to the functioning of the world, they quote Ager (1973, p. 114):³⁶ “The history of any one part of the earth, like the life of a soldier, consists of long periods of boredom and short periods of terror”.³⁷

Raup (1966) exploits the evolution of complex life through the study of the geometry of shells that grow in spirals. He postulates that the format of shells can be expressed by geometric parameters and simulates computationally the formats that could exist theoretically. The comparison of the simulation with real species shows that the known species in nature are not randomly distributed within all possible formats. On the contrary, they concentrate on discrete regions of the distribution. Raup suggests that there are rational explanations for this distribution which could be potentially associated to the functional characteristics of the shells. His work is an example of the evolutionary exploitation of the morphospace and of the use of computational simulation to study the growth and form of organisms.

Maynard Smith (1974) applies game theory to analyze the evolution of patterns of behavior in conflicts among animals and establishes the concept of evolutionary stable strategy (ESS). A strategy I is considered ESS, if the expected utility of a player using strategy I against himself is higher than the expected utility of strategy J applied against strategy I .

$$E_I(I) > E_I(J) \tag{10}$$

where E indicates the expectancy of the utility of the strategy in parenthesis against the other strategy.

In a population composed of individuals who adopt strategy I , initially rare mutant individuals, adopting a strategy J different from I would not increase in frequency so that the population would remain stable even with the occurrence of mutation and natural selection. The concept of ESS can be considered a particular case of Nash equilibrium.

36. Ager, D.V. *The nature of the stratigraphical record*. New York: John Wiley, 1973. p. 114.

37. Note that the dispute among gradualists and punctualists persists today.

The concept was also absorbed by other fields of science such as Economics and Political Science. The ESS can be said to have contributed greatly to the study of evolution.³⁸

Holland and Reitman (1977) propose the use of simple learning genetic algorithms that can: *i*) model the history of successful agents with the roles of classifiers; and *ii*) update themselves into more efficient algorithms when necessary. In fact, the algorithm tends to favor generic agents who can consistently make more precise predictions. Besides, the genetic algorithm allows for the construction of new agents from the combination of parts chosen by performance rather than randomly, maintaining the efficiency of the results at each interaction.

Two models are tested. The first confirms the magnitude of the power of the result reached with the genetic algorithm when compared to the traditional one. The second model compares the first genetic algorithm with another genetic algorithm with learning capacity. This second model also results in significant gains. Models based on the sending and receiving of messages among the agents are described as computationally complete, which means that they can be tested independently of the chosen platform.³⁹

Waddington (1942) writes about the dispute at his time between naturalists and geneticists highlighting the prominence of genetics as a theory based on evidence. He understands the discontentment of naturalists as he knows that assuming natural selection to be the random mechanism that is the sole responsible for observed adaptation seems insufficient. And this is exactly the objective of Waddington's paper: to demonstrate the processes that facilitate the understanding of how geneticists are right when evocating the genes as the ones responsible for a coordinated evolution in the duet environment-organisms.

The central argument developed by Waddington is that adaptations may occur as answers to the environment, but that the genetic factor finally surpasses the initial environment effect. In fact, adaptation is only possible because the answer itself is already genetically programmed (although dormant)⁴⁰ and is only "activated" by the environment.

38. Available at: <<http://goo.gl/2QDZ1t>>.

39. For a deeper view see Booker, Goldberg and Holland (1989).

40. Our interpretation.

The second argument is the canalization – given as “a measure of the ability of a population to produce the same phenotype regardless of the variability of its environment or genotype”.⁴¹ In other words, Waddington defines canalization as: “[developmental reactions that] are adjusted so as to bring about one definite end-result regardless of minor variations in conditions” (Waddington, 1942, p. 563). Further on, Waddington makes clear the idea that a channel is a conveyor of alternatives that are activated or not depending on the influence of the environment or other correlated processes that work as switches.

The occurrence of an adaptive response to an environmental stimulus depends on the selection of a suitable genetically controlled reactivity in the organism. If it is an advantage... then the reactivity will become canalized, again under the influence of natural selection (Waddington, 1942, p. 565).

8 CRITICISMS⁴²

Israel (2005) defies the notion that the contribution of the science of complexity is innovative and questions if it can even be considered a science at all. He starts his critic stating that there is no uncontested definition in the literature.⁴³ In fact, von Neumann (1966) uses the terms complex and complicated indistinctly, as mentioned in the introduction.

Another criticism from Israel is over the supposedly larger view of phenomena – supported by the relevance given to the dynamics of the systems. Israel claims that such focus on intertemporal interactions and on states of phenomena are in fact more restrictive than the traditional paradigm.

Essentially, the objective of discussing the “science of complexity” clashes with a two-fold difficulty: the absence of a rigorous definition of complexity and thus the impossibility of arriving at a precise definition of the object and methods of this “science”; the high degree of articulation of the historical pathways leading to the topics of complexity, that is, to something that is ill-defined, “a style of theorizing” rather than a specific subject matter (Israel, 2005, p. 481).

41. Definition available at: <<http://goo.gl/e42Tyr>>.

42. Again, the authors reaffirm that the choice of the authors to be included in this review to be presented as classics were those selected by the Santa Fe and listed as background papers.

43. In fact, this is not disputed by those working with complexity. See the interviews made by Melanie Mitchell in the online course *Introduction to complexity*, available at: <<http://goo.gl/Xjhfca>>.

Another criticism has to do with the allegation made by the science of complexity that it is not possible to keep the scientific reductionism paradigm, given that the sum of the parts is not sufficient to explain the whole of phenomenon. The whole would be “more” than the sum of the parts. In other words, the complex approach advocates that there are “new or unexpected” mechanisms in the interactions among the parts that only emerge when observing the whole set of interactions among individuals.

Still against the concept that the whole is “more than” the sum of the parts, Israel argues that the usual definition of “the whole as the sum of the parts” does not imply the absence of emerging properties and that, in fact, the properties are themselves effects that occur as a result of the action of the parts (which include their interactions). Thus, this synergy among the parts and their emerging properties are already included in recent mathematical and computational development. Anyway, it would have been contradictory to define an object as the sum of the parts and at the same time to say that an emerging property exists because the system behaves differently than the sum of its parts.⁴⁴

Another argument made by Israel is that there is no dichotomy between linear and nonlinear mathematics. Israel ponders that linearization is central as a problem solving technique in mathematics but he reinforces that the tools available are not enough to explain phenomena as complex as social ones.

Israel challenges the notion that the science of complexity is innovative and claims that the “new” science of complexity only aggregates fragmented advances which are defined within three great lines of thought: information theory – in which there is mathematical precision in its definition; chaos theory and game theory. He argues that if the concepts are already present in these lines of science, putting them together does not make a “new” science. Israel lists the concepts that are already present in those three lines of research.

1. Information theory includes the aspects related to cyber studies, feedback analysis, self-organizing systems and homeostasis.

44. Although we recognize the epistemological criticism made by Israel, we highlight that the emphasis of the science of complexity is exactly that, i.e., that the structural behavior can be different from the behavior of the parts. Describing the parts, then, is not enough to characterize the system.

2. Chaos theory and dynamic systems have characteristics that are typically reductionist. Yet, Israel says that chaos can be precisely defined, but that does not make the definition useful, only interesting. Israel also talks about the impossibility of chaos in models as it is “only” a result of a mathematical fact.
3. Game theory and genetic algorithms are based on the theory of evolution from biology and by themselves do not constitute a “new science”.

Another key element of Israel’s criticism is the difficulty of complexity science to handle two relevant principles of scientific analysis as we know today: objectivity and previsibility.

Israel defends that previsibility of phenomena is possible but that it would be too costly to model such cases. He explains himself using the known difficulties to predict the weather using mathematics. In sum, he refutes the conclusion that ‘indeterminism’ may result from “deterministic” processes.⁴⁵ In reality, he reaffirms the validity of the Laplacian thought. Laplace defended the view that previsibility is possible only in an epistemological sense. He says that if knowledge of all particles, all initial conditions, and all rules was/were available, prediction would be possible. However, based on the concept of model that is necessarily smaller than the object being modeled, Israel (2005) states that any representation is unequivocally imprecise, by construction.⁴⁶ However, he ponders that one cannot expect absolute prevision, even given the knowledge of natural laws.

Israel argues further that even those that criticize the problems of reductionism are themselves reductionists as the simple search for explanations imply in simplification and thus reduction. He says that researchers are replacing a reductionism from physics and moving into a reductionism from biology that favors the rational quest for goals. He says that in this view, natural selection could not be used in all cases, but only in particular ones.

45. Moreover, says Israel, even if there were deterministic chaos, then it would have been impossible to make predictions. All in all, the theory of dynamic systems has moved further and this concept (of not having a deterministic chaos) is no longer disputed. Finally, it is relevant to say that chaos theory predicts intervals and parameters where there is chaos. In many systems, its evolution leads to homogeneous patterns, stable; or to recurrent attractors, also stable. See Strogatz (2001).

46. In fact, this is discussed on the paper by Rosenblueth and Wiener (1945).

Still discussing evolution, Israel concedes the real and old difficulty of probabilistic and deterministic tools to handle consciousness and agents that actively pursue goals (such as citizens). The real distinction would be between teleological approaches and causal ones with the former being more general. Yet, from an epistemological perspective there is no real difference between the approaches.

In his conclusion, Israel reaffirms the characteristic of complexity science as “an updated version of reductionism” (Israel, 2005, p. 505), justifying that the formal mathematics remains theoretically central. He confirms that the science makes use of the same tools, objectives and functions that it criticizes. He does not reject the advances proposed by the studies of complexity; he only reaffirms that the arguments do not support the constitution of a new science.

9 CONCLUDING REMARKS

Israel (2005) criticizes the argument that the concepts presented in this paper together constitute a science of complexity. In fact, he may just be right. Anyway, not constituting a science does not change the fact that for a number of systems classified as complex, such as the economy, society, physics or biology, the concepts presented bring analytical power and contribute to the advance of disciplines and the cross-fertilization among disciplines.

Israel’s suggestion is that authors who are interested in the objects and phenomena characterized as complex should focus on the subjective analysis and historicity of the facts. We disagree. Whether they constitute a science or not, the notions presented contribute fundamentally to the scientific approach in at least two manners.

On one hand the concepts bring formalism and theoretical scope for the discussions in areas of the science that are less used to them. On the other hand, the concepts effectively handle elements that even in the so-called hard sciences are routinely ignored given their frequent mathematical intractability, that is: *i*) the dynamics of the events and their strong crossed causality (interaction) that includes feedback elements; *ii*) the randomness of events that may be relevant and together lead to rupture points and structural change; and *iii*) self-organization without central control.

Such elements should be considered by the scientist every time he or she analyzes phenomena that have heterogeneous agents interacting with each other, and with the environment in a dynamic and nonlinear manner, in systems that have self-organization and differentiated characteristics for different scales; and that have the ability to evolve and adapt. In sum, one should proceed with those concepts in mind when studying systems which are clearly observable and frequent in nature and in society. Such scientist should also model the system at hand meaning that the essence of the work should consist in finding the key elements driving the phenomena being investigated.

This paper aimed at bringing together initial fragmented contributions, definitions and concepts that can contribute to the scientifically rigorous analysis of complex phenomena. From the point of view of the authors who are affiliated to an institute focused on studies of the economy, the society, the cities, and the ecology, the concepts described in this text seem relevant to better model (and simulate) the environment of applied public policies.

REFERENCES

- ANDERSON, P. W. More is different. **Science**, Washington, v. 177, n. 4.047, p. 393-396, 8 Apr. 1972.
- BENNETT, C. H. Logical reversibility of computation. **IBM journal of research and development**, v. 17, n. 6, p. 525-532, 1973.
- BOOKER, L. B.; GOLDBERG, D. E.; HOLLAND, J. H. Classifier systems and genetic algorithms. **Artificial intelligence**, v. 40, n. 1, p. 235-282, 1989.
- CONANT, R. C.; ASHBY, W. R. Every good regulator of a system must be a model of that system. **International journal of systems science**, v. 1, n. 2, p. 89-97, 1970.
- DERRIDA, B.; GARDNER, E.; ZIPPELIUS, A. An exactly solvable asymmetric neural network model. **Europhysics letters**, Bristol, v. 4, n. 2, p. 167, July 1987.
- GELL-MANN, M.; LLOYD, S. Effective complexity. *In*: GELL-MANN, M.; TSALLIS, C. **Nonextensive entropy**: interdisciplinary applications. Oxford: SFI, 2004. p. 387-398.
- GOULD, S. J.; ELDREDGE, N. Punctuated equilibria: the tempo and mode of evolution reconsidered. **Paleobiology**, v. 3, n. 2, p. 115-151, 1977.

HOLLAND, J.; REITMAN, J. Cognitive systems based on adaptive algorithms. **SIGART newsletter**, New York, n. 63, p. 49, 1977.

HOPFIELD, J. J. Neural networks and physical systems with emergent collective computational abilities. **Proceedings of the national academy of sciences**, Washington, v. 79, n. 8, p. 2.554-2.558, 1982.

ISRAEL, G. The science of complexity: epistemological problems and perspectives. **Science in context**, Cambridge, v. 18, n. 3, p. 479-509, 2005.

KOLMOGOROV, A. N. Three approaches to the quantitative definition of information. **Problems of information transmission**, v. 1, n. 1, p. 3-11, 1965.

LANDAUER, R. Irreversibility and heat generation in the computing process. **IBM journal of research and development**, v. 5, n. 3, p. 183-191, July 1961.

LANGTON, C. G. Studying artificial life with cellular automata. **Physica D: nonlinear phenomena**, v. 22, n. 1, p. 120-149, 1986.

MAYNARD SMITH, J. The theory of games and the evolution of animal conflicts. **Journal of theoretical biology**, v. 47, n. 1, p. 209-221, 1974.

MILLER, J. H.; PAGE, S. E. **Complex adaptive systems**. New Jersey: Princeton University Press, 2007.

MINSKY, M. Steps toward artificial intelligence. **Proceedings of the IRE**, v. 49, n. 1, p. 1-65, Jan. 1961.

MITCHELL, M. **Complexity: a guided tour**. Oxford; New York: Oxford University Press, 2011.

PAGE, S. E. **The difference: how the power of diversity creates better groups, firms, schools, and societies**. 1. ed. New Jersey: Princeton University Press, 2008.

_____. **The model thinker**. Ann Arbor: Center for the Study of Complex Systems, 2012. Disponível em: <<http://goo.gl/NuOxI7>>. Acesso em: 21 dez. 2012.

RAUP, D. M. Geometric analysis of shell coiling general problems. **Journal of paleontology**, Connecticut, v. 40, n. 5, p. 1.178-1.190, 1966.

ROSENBLUETH, A.; WIENER, N. The role of models in science. **Philosophy of science**, Chicago, v. 12, n. 4, p. 316-321, 1945.

SHANNON, C. E. A mathematical theory of communication. **The bell system technical journal**, v. 27, p. 379-423; p. 623-656, 1948.

SIMON, H. A. The organization of complex systems. In: PATTEE, H. H. (Ed.). **Hierarchy theory: the challenge of complex systems**. New York: George Braziller, 1973. p. 1-27.

STROGATZ, S. H. **Nonlinear dynamics and chaos**: with applications to physics, biology, chemistry, and engineering. 1st ed. Boulder: Westview Press, 2001.

TOBLER, W. R. A computer movie simulating urban growth in the Detroit region. **Economic geography**, Worcester, v. 46, p. 234-240, 1970.

TURING, A. M. The chemical basis of morphogenesis. Philosophical transactions of the Royal Society of London. **Biological sciences**, v. 237, n. 641, p. 37-72, Aug. 1952.

VON NEUMANN, J. The role of high and of extremely high complication. *In*: _____. **Theory of self-reproducing automata**. Urbana: University of Illinois Press, 1966. p. 64-87.

WADDINGTON, C. H. Canalization of development and the inheritance of acquired characters. **Nature**, v. 150, n. 3.811, p. 563-565, 1942.

WOLFRAM, S. Universality and complexity in cellular automata. **Physica D: nonlinear phenomena**, v. 10, n. 1, p. 1-35, 1984.

_____. **A new kind of science**. Champaign: Wolfram Media, 2002.

PUBLISHING DEPARTMENT

Coordination

Cláudio Passos de Oliveira

Supervision

Everson da Silva Moura

Reginaldo da Silva Domingos

Translation

Bernardo Alves Furtado

Patrícia Alessandra Morita Sakowski

Typesetting

Bernar José Vieira

Cristiano Ferreira de Araújo

Daniella Silva Nogueira

Danilo Leite de Macedo Tavares

Diego André Souza Santos

Jeovah Herculano Szervinsk Junior

Leonardo Hideki Higa

Cover design

Luís Cláudio Cardoso da Silva

Graphic design

Renato Rodrigues Buenos

*The manuscripts in languages other than Portuguese
published herein have not been proofread.*

Ipea Bookstore

SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo

70076-900 – Brasília – DF

Brazil

Tel.: + 55 (61) 3315 5336

E-mail: livraria@ipea.gov.br

Ipea's mission

Enhance public policies that are essential to Brazilian development by producing and disseminating knowledge and by advising the state in its strategic decisions.

