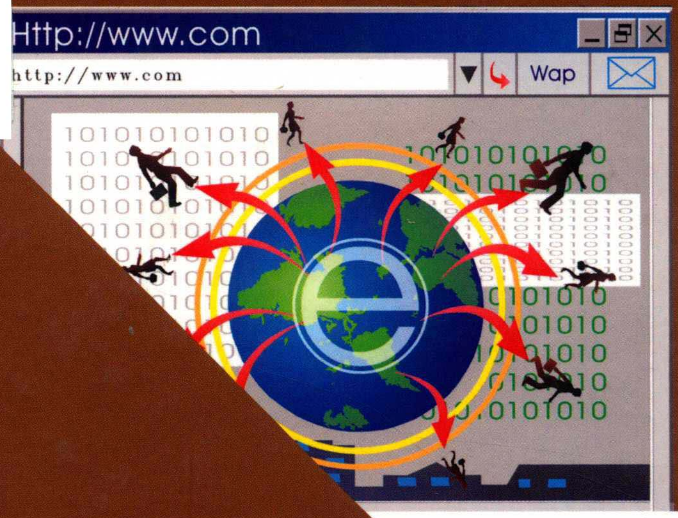


World Scientific Series in Information Studies – **Vol. 6**

INFORMATION AND COMPLEXITY

Edited by

Mark Burgin

Cristian S Calude

World Scientific

World Scientific Series in Information Studies – Vol. 6

INFORMATION AND COMPLEXITY

The book is a collection of papers of experts in the fields of information and complexity. Information is a basic structure of the world, while complexity is a fundamental property of systems and processes. There are intrinsic relations between information and complexity.

The research in information theory, the theory of complexity and their interrelations is very active. The boo n, complexity and their relations repr d studies and achievements in this are

The goal of the book is to present – mathematical, informational, philosophical, methodological, etc.

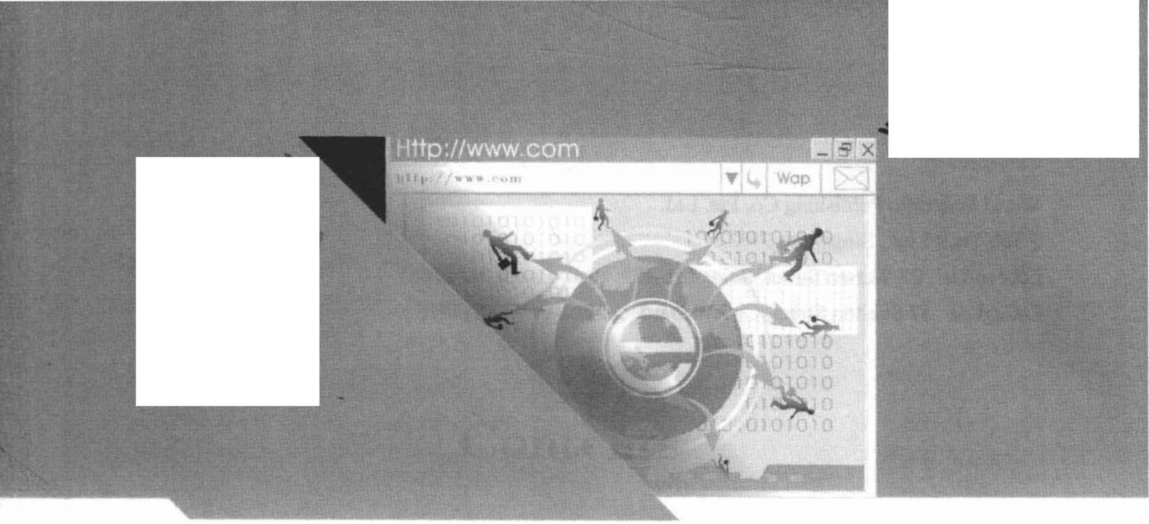
World Scientific
www.worldscientific.com
10017 hc

ISBN 978-981-3109-02-5
9 789813 109025

Burgin • Calude

Vol. 6

INFORMATION AND CO



World Scientific Series in Information Studies — **Vol. 6**

INFORMATION AND COMPLEXITY

Edited by

Mark Burgin
University of California, Los Angeles, USA

Cristian S Calude
University of Auckland, New Zealand

World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI • TOKYO

Published by

World Scientific Publishing Co. Pte. Ltd.
5 Toh Tuck Link, Singapore 596224
USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601
UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Cataloging-in-Publication Data
Names: Burgin, M. S. (Mark Semenovich) | Calude, Cristian, 1952–
Title: Information and complexity / [edited by] Mark Burgin (UCLA), Cristian S. Calude (University of Auckland, New Zealand).
Description: New Jersey : World Scientific, 2016. | Series: World scientific series in information studies ; volume 6
Identifiers: LCCN 2016023758 | ISBN 9789813109025 (hc : alk. paper)
Subjects: LCSH: Information theory. | Computational complexity.
Classification: LCC Q360 .I5275 2016 | DDC 003/.54--dc23
LC record available at https://lccn.loc.gov/2016023758

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Desk Editors: V. Vishnu Mohan/Tan Rok Ting

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

INFORMATION AND COMPLEXITY

World Scientific Series in Information Studies
(ISSN: 1793-7876)

Vol. 1 Theory of Information — Fundamentality, Diversity and Unification
by Mark Burgin

Vol. 2 Information and Computation — Essays on Scientific and Philosophical Understanding of Foundations of Information and Computation
edited by Gordana Dodig-Crnkovic & Mark Burgin

Vol. 3 Emergent Information — A Unified Theory of Information Framework
by Wolfgang Hofkirchner

Vol. 4 An Information Approach to Mitochondrial Dysfunction: Extending Swerdlow's Hypothesis
by Rodrick Wallace

Vol. 5 Theory of Knowledge: Structures and Processes
by Mark Burgin

Vol. 6 Information and Complexity
edited by Mark Burgin & Cristian S Calude

Vol. 7 Information Theory Models of Instabilities in Critical Systems
by Rodrick Wallace

Vol. 8 The Future Information Society: Social and Technological Problems
edited by Wolfgang Hofkirchner & Mark Burgin

Preface

The Interplay Between Information and Complexity

Although today information and complexity have become popular and important notions in science and society, experts agree that it is hard (if not impossible, as some think) to define them (Capurro *et al.*, 1999; Capurro and Hjorland, 2003; Gell-Mann and Lloyd, 1996; Burgin, 2010). At the same time, a diversity of definitions have been suggested for both terms. For instance, 130 definitions of data, information and knowledge formulated by 45 scholars are collected in Zins (2007), while a variety of approaches to define information appears in the book of Burgin (2010) and in the encyclopedia article of Bates (2010).

Information is related to everything and everything is related to information. Why is it so difficult to define information? There are many factors contributing, including the diversity of information types, contradictory usages of the term information and the confusion between information and data or knowledge, on the one hand, and information carriers (such as messages) and information measures (such as Shannon's entropy), on the other hand. An endeavor to create a general theory of information based on a system of postulates (principles) is presented in Burgin (2010) and in the mathematical form in Burgin (2010a).

Similarly, there is no general definition of complexity. The "symptom" or "mark" of complexity is generally associated with some action. The complexity of a system or process is related to some ways to describe it, to build it, to control it, and so on. High complexity appears in systems composed of many parts, which interact with each other in multiple ways. However, when the action is, for example, the possibility to give an algorithmic description of a system, the complexity can be low because a simple algorithm can describe the system.

Complexity can be dynamic, like time or space complexity, or descriptive, as size complexity. Complexity studies appear in a multitude of areas and domains including, but not restricted to:

(i) Computer science – computational complexity (Trakhtenbrot, 1956; Rabin, 1959), algorithmic information theory via Kolmogorov complexity (Solomonoff, 1964; Kolmogorov, 1965; Chaitin, 1977) and its generalizations), Boolean circuit complexity (Savage, 1976), structural complexity (Balcazar *et al.*, 1988), communication complexity (Yao, 1979);
(ii) Physics – statistical mechanics entropy (Lavenda, 2010);
(iii) Mathematics – Krohn–Rhodes complexity (Kambites, 2007);
(iv) Complex systems (Grassberger, 1990);
(v) Cognitive sciences (Isaac *et al.*, 2014).

The fact that there are many forms (definitions) of and approaches to complexity is not a weakness nor a sign of lack of maturity. Indeed, neither form of complexity is "better" than the other; they are just targeting different aspects. Studying the algorithmic complexity of the genetic code may help understanding the redundancy of the genetic code, its role in inheritance and evolution and could be used in genetic engineering problems. Studying the computational complexity of the process of translating the genetic code into anatomical details of a living being is different and important for other reasons.

The most popular theory of information is Shannon's theory (Shannon, 1949), which studies the transmission, processing, utilization, and extraction of information by a purely syntactic method, devoid of any meaning. Entropy, a key measure in this theory, quantifies the uncertainty involved in predicting the value of a random variable. The theory has the broadest range of applications, from engineering and data analysis to biology and linguistics.

Algorithmic information theory, based on the concept of Kolmogorov complexity or more generally, on the algorithmic complexity of objects, combines the theory of algorithms with Shannon's information theory by measuring the size of algorithmic descriptions of (finite or infinite) objects (Calude, 2002; Li and Vitanyi, 1997). The theory has been applied to a wide range of areas including theory of computation, physics, biology, economics, combinatorics, inductive reasoning, machine learning and cognitive sciences.

Information is intimately related to energy. According to John Wheeler, every physical quantity derives its ultimate significance and meaning from information. He aptly expressed it by "It from Bit", where It stands for things, while Bit refers to the most popular information unit (Wheeler, 1990). The principles of a general theory of information developed in Burgin (2003; 2010) demonstrate that energy is a kind of information in a broad sense.

There is also a very strong relation between complexity and information. As a simple observation, we note that measuring complexity is a way to understand information, while obtaining information helps to decrease the complexity. The duality between complexity and information plays an important role in impossibility results, which reflect infinite complexity and the inaccessibility of exact information. The impossibility of constructing a *perpetuum mobile* led to the notion of energy and to the development of thermodynamics, the impossibility of traveling faster than light was a starting point of relativity theory, and the impossibility of measuring the speed and position of a particle simultaneously is fundamental in quantum mechanics. The impossibility results regarding complexity and information – for example, the impossibility of solving the Halting Problem for Turing machines or, more generally, the undecidability results in computability theory (Cutland, 1980) – have far reaching consequences. The Halting Problem for Turing machines is the problem of deciding in *finite time* whether an arbitrary Turing machine stops or not on a given input. It is interesting to examine the reason of its undecidability from an informational perspective. The restriction to "finite time" is crucial and the problem can or cannot be solved depending on the tools processing the given information: Turing machine and input. If we use Turing machines to solve the problem, the given information remains inaccessible, hence the Halting Problem for Turing machines is undecidable. However, the more powerful inductive Turing machines can retrieve and "extract" this information to solve the Halting Problem for Turing machines (Burgin 2001). Furthermore, one can also "approximately" solve the Halting Problem with arbitrary precision (Calude and Dumitrescu 2015).

In this volume, leading researchers present recent results in a variety of areas related to information and complexity. The contributions are divided into three groups comprising three parts of the book: (1) classical information and complexity, (2) quantum information and complexity, and (3) complexity and information applications. Inside each group, the authors are ordered alphabetically.

The chapter "The 'Paradox' of Computability and a Recursive Relative Version of the Busy Beaver Function" by Felipe S. Abrahão studies relative properties of computability in the context of metabiology and total Turing machines. Metabiology is a field of theoretical computer science that studies general principles of biological relations at a meta-level using methods from evolutionary biology and algorithmic information theory (Chaitin, 2014). The author proves that a Busy Beaver function introduced by G. Chaitin behaves to total Turing machines in the "same" way as the original Busy Beaver function behaves to all Turing machines. This is termed a "paradox" of computability *a la* Löwenheim-Skolem (Löwenheim, 1915; Skolem, 1920) – a function that is computable when "seen from the outside" the subsystem, but incomputable when "seen from within" the same subsystem. This result is aimed at modeling biological phenomena. From the perspective of information, Abrahão illustrates how information inaccessible within a system becomes accessible in a larger system. From the viewpoint of complexity, this chapter shows how complexity of a problem is decreased from infinity to a finite quantity.

Shannon entropy has become the most popular measure of information. In his chapter "Inductive complexity and Shannon entropy", Mark Burgin studies another measure of information, called inductive complexity. This complexity is similar to Kolmogorov complexity: instead of Turing machines it uses inductive Turing machines, which are super-recursive algorithms because they can compute much more than recursive algorithms (Burgin, 2001). The main goal is to compare inductive complexity as a measure of information with Shannon entropy. It is proved that the average of the prefix inductive complexity $\mathrm{IK}(x)$ with respect to an inductively computable probability distribution f is equal to Shannon entropy up to some additive constant that depends on f. A similar inequality was proved for recursively computable probability distributions and the conventional (Kolmogorov) prefix complexity (Grunwald and Vitanyi, 2004; Muchnik and Vereshchagin, 2006). The result proved in this chapter has the following advantages:

1. There are more inductively computable probability distributions than recursively computable probability distributions.
2. For infinitely many functions f, the estimate $\mathrm{IK}(f)$ is more exact than the estimate $\mathrm{K}(f)$ obtained prefix complexity because the value of $\mathrm{IK}(f)$ is smaller than the value of $\mathrm{K}(f)$ (Burgin, 2005).

Estimating relations between inductive complexity and Shannon entropy, Burgin uses only inductive Turing machines of the first order. However, there is an advanced hierarchy of inductive Turing machines (Burgin, 2005). Thus, it would be interesting to find relations between Shannon entropy and inductive complexity based on inductive Turing machines of higher orders.

One of the most important ways to understand information and complexity is through the axiomatic method. Axiomatic approach to complexity started with Blum, who defined the size of a machine as a direct complexity measure by two simple axioms (see (Blum, 1967) and the presentation in Calude (1988)). To develop axiomatic foundations for Kolmogorov complexity and algorithmic information theory, Burgin (1982) enhanced and expanded Blum's axioms to a system of five axioms building static dual complexity measures, which encompassed all known forms of algorithmic complexity. In his chapter "Blum's and Burgin's Axioms, Complexity, and Randomness", Cezar Câmpeanu explores relations between complexity, randomness and information in the context of the axiomatic theory of dual static complexity measures. The author uses this framework to reveal the important connections between information and complexity, on the one hand, and algorithmic randomness for infinite sequences, on the other hand.

In his chapter "Planckian Information (I_P): A Measure of the Order in Complex Systems", Sungchul Ji constructs a new information measure called Planckian information. He shows that Planckian information is a better measure of order in organized complex systems than either the Boltzmann-Gibbs entropy S or Shannon entropy H. This statement is supported with examples from atomic physics, protein folding, single-molecule enzymology, whole-cell mRNA metabolism, T-cell receptor variable region diversity, fMRI (Functional Magnetic Resonance Imaging), the quantitative study of words and texts, econophysics, and astrophysics.

In algorithmic information theory, complexity and randomness are studied for relatively simple objects such as sequences of symbols from a given alphabet. In the chapter "On Algorithmically Random Universal Algebras", Bakhadyr Khoussainov extends this approach to mathematical systems with additional structures, such as universal algebras, and develops an abstract mathematical framework for answering the following question: What is an algorithmically random infinite universal algebra? The proposed definition of algorithmic randomness for infinite universal algebras is based on Martin-Löf tests of randomness and its adequacy is justified by explication of some natural properties of random algebras.

In his chapter "Structural and Quantitative Characteristics of Complexity in Terms of Information", Marcin J. Schroeder studies issues related to overcoming or controlling complexity seen in the entire intellectual development of humanity, although, historically, the emphasis was on its opposition – simplicity. First, the author reviews various conceptualizations of complexity in the European philosophical and scientific tradition from Pre-Socratic thinkers of Antiquity to the present time. Special focus is given on the contributions of Aristotle whose works strongly influenced the view on reality of his followers, as well as of his adversaries. Then, this historical perspective is used for a critical summary of the ways complexity is studied today and for an attempt to provide a unified conceptual framework for the various forms of complexity considered in the past. To this purpose, the author uses the general concept of information developed in Schroeder (2005) as an identification of a variety and various concepts characterizing it. Finally, a quantitative description in the form of two related measures of complexity and information integration is presented.

In the chapter "Multiscale Information Theory for Complex Systems: Theory and Applications", Blake Stacey, Benjamin Allen and Yaneer Bar-Yam construct a mathematical formalism based on information theory to make precise the intuition that a complex system exhibits structure at multiple scales. They correctly assume that science in general and the science of complex systems, in particular, require a general way to understand and represent structure. Even in mathematics, the concept of structure is still in the process of development and a comprehensive mathematical representation of this concept is elaborated in the general theory of structures (Burgin, 2012).

This article treats structure as the totality of relationships between a system's components, which is according to the general theory of structures, a first-order inner structure of a system, and uses information theory to quantify these relationships. To this aim Stacey Allen and Bar-Yam create an axiomatic system for characterizing an information measure by developing quantitative indices – the complexity profile (CP) and the marginal utility of information (MUI) – which describe the system structure.

A complex system exhibits structure at many scales and levels of organization. For instance, one can study human beings at any magnification, from the molecular level to the societal. Thus, the concept of a scale becomes crucial and the authors of this chapter use the mathematical formalism developed in their previous work (Allen *et al.*, 2014) to operate with it. Often, *scale* is interpreted in terms of length or of time. However, in the context of system theory developed in this chapter, a scale reflects the number of components in the system. Besides, in the context of multiscale physics and scalable topology, a scale C of a system R is the theoretical (mathematical) structure used to model R (Burgin, 2006). This chapter also demonstrates the applications of the proposed theory to evolutionary biology, economics and finite geometry.

Kolmogorov complexity is one of the most important measures of information and complexity. In his chapter "Bounds on the Kolmogorov complexity function", Ludwig Staiger presents a detailed analysis of the Kolmogorov complexity function examining its bounds for maximally complex infinite sequences. Lower and upper bounds on the Kolmogorov complexity function are closely related to algorithmic randomness, in particular to partial randomness. The performed analysis provides interesting connections between information, complexity, Hausdorff dimension, fractal geometry and randomness.

The main goal of quantum information science is to use quantum phenomena to improve the performance of information processing systems. The current theoretical progress led in some cases to commercial applications. As Marco Lanzagorta and Jeffrey Uhlmann note in their chapter "Quantum Computational Complexity in Curved Spacetime", the "optimism about their potential advantage over classical alternatives is invariably founded on the theoretical analysis that implicitly assumes a flat spacetime operating environment, i.e., that their operations will not be influenced by local gravitation". However, gravity affects quantum information, hence, its role requires a careful analysis. This paper develops a new quantum computational complexity in which gravity is no longer ignored.

In her chapter "A Silk Road from Leibniz to Quantum Information", Rossella Lupacchini brings about interesting relations between the monads of Leibniz and qubits of contemporary quantum information theory exploring the meaning of the principles of continuity and distinguishability as well as of the notions of causality and correlation. From Leibniz's point of view, the ontological separation of *nature* and *thought*, which pervades the Western philosophy, violates the principle of continuity. Lupacchini compares Leibniz's pre-established harmony and the harmony emerging from Chinese "correlative thinking" in order to contrast the *principle of sufficient reason* with the Chinese *principle of connectedness*. According to Leibniz's metaphysics, there must be a reason to cause the harmony of monads, while in the ancient Chinese thought the fundamental harmony of the universe has no reason, it is a harmony of resonant possibilities. In this perspective, the divide between Leibniz's causal determinism and quantum randomness turns into a *coincidentia oppositorum*.

It might be helpful to frame Leibniz's *Monadology* in millennia of philosophical search for a "fundamental principle", unit or essence from which all things originate. Plato, for instance, conjectured that all things are drawn from triangles. Democritus asserted that everything is made from atoms. In his turn, Leibniz conceived monads as elementary constituents for all things. As a simple substance and a unity of perceptions, however, each monad encapsulates all information of the universe. Indeed, monads are viewed as "incorporeal automata". Now, as elements of things, the monads are "the true atoms of nature", but as they have neither extension nor shape, they are unobservable. In mathematical terms, the essential character of monads, which is continuous change, is captured by a unitary transformation demonstrating that it is intrinsically formed as a fundamental triad (Burgin, 2012). Thus, Lupacchini regards Leibniz's monads as remote ancestors of qubits, the building blocks of quantum information theory.

For the classical mind, quantum mechanics is boggling enough. Nevertheless, in the chapter "Generalized Event Structures and Probabilities", Karl Svozil studies an even more sophisticated world where basic structures can transcend the quantum domain. In it, predictions and probabilities are neither classical nor quantum, but are subject to sub-classicality, that is, the additivity of probabilities for mutually exclusive, co-measurable observables, is formalized by admissibility rules and frame functions. Starting with the Specker's oracle, Svozil develops an information-theoretic framework for analyzing a strange, but logically not impossible, universe through a calculus with generalized event structures and their probabilities.

In the chapter "An upper bound on the asymptotic complexity of global optimization of smooth univariate functions", James M. Calvin studies complexity of adaptive optimization algorithms using real-number model of computations. The main characteristic of adaptive algorithms is that to make the next step of computation, they utilize information gained in the previous steps. Efficient utilization of such information allows decreasing complexity of algorithms.

Set theory has no informational tools to distinguish between the "size" of the set of natural numbers and the "size" of the set of all even natural numbers. Intuitively, the size of the later set is one half of the size of the former one. Can this statement be rigorously proved? The answer is affirmative. A formal method introduced by Sergeyev uses the resource-based informational new language of Grossone to deal with infinities and infinitesimals in a computational way (Sergeyev, 2009). Another approach to the rigorous distinction between the "sizes" of the set of all natural numbers and the set of all even natural numbers is based on utilization of hypermeasures. Note that the measure of set is a kind of its complexity, while measuring provides information about this complexity. In their chapter "Cellular Automata and Grossone Computations", Louis D'Alotto and Yaroslav Sergeyev use the language of Grossone to perform computations on cellular automata. This allows measuring the quantity of information that can pass through a central viewing window under a cellular automaton rule.

In the chapter "Cognition and Complexity", Yuri Manin analyses scientific cognition as an information process by means of Kolmogorov complexity. The author shows that many scientific theories, from Ptolemy's epicycles to the Standard Model of elementary particles, consist of two distinct parts: the part with the relatively small Kolmogorov complexity ("laws", "basic equations", "periodic table", "natural selection, genotypes, mutations") and another part, of indefinitely large Kolmogorov complexity ("initial and boundary conditions", "phenotypes", "populations"). The data for the latter part are obtained by planned observations, focused experiments, and are afterwards collected in growing databases (formerly known as "books", "tables", "encyclopedias", etc.). The former part is meant to "make sense" of the latter one.

In his chapter "Informational Perspective on QBism and the Origins of Life", Koichiro Matsuno studies the interpretation of quantum mechanics called Quantum Bayesianism or QBism with the emphasis on the role of information in understanding quantum processes. Then Matsuno uses his approach as a tool for exploration of the origins of life as emergence of an observer of quantum processes considered in quantum mechanics.

Bibliography

Allen, B., Stacey, B. C. and Bar-Yam, Y. (2014) An information-theoretic formalism for multiscale structure in complex systems, arXiv:1409.4708 [cond-mat.stat-mech].

Balcazar, J.L., Diaz, J., and Gabarro, J. (1988) *Structural Complexity*, Springer-Verlag, Berlin/Heidelberg/ New York.

Bates, M. J. (2010) Information, in *Encyclopedia of Library and Information Sciences*, Third Edition, v. 1, pp. 2347-2360.

Blum, M. (1967) On the size of machines, *Information and Control*, v. 11, pp. 257-265.

Burgin, M. (2001) How We Know What Technology Can Do, *Communications of the ACM*, v. 44, No. 11, pp. 82-88.