Studies in Applied Philosophy,
Epistemology and Rational Ethics

SAPERE

Gordana Dodig-Crnkovic
Raffaela Giovagnoli *Editors*

Representation and Reality in Humans, Other Living Organisms and Intelligent Machines

Springer

Studies in Applied Philosophy, Epistemology and Rational Ethics

Volume 28

About this Series

Studies in Applied Philosophy, Epistemology and Rational Ethics (SAPERE) publishes new developments and advances in all the fields of philosophy, epistemology, and ethics, bringing them together with a cluster of scientific disciplines and technological outcomes: from computer science to life sciences, from economics, law, and education to engineering, logic, and mathematics, from medicine to physics, human sciences, and politics. It aims at covering all the challenging philosophical and ethical themes of contemporary society, making them appropriately applicable to contemporary theoretical, methodological, and practical problems, impasses, controversies, and conflicts. The series includes monographs, lecture notes, selected contributions from specialized conferences and workshops as well as selected Ph.D. theses.

Advisory Board

More information about this series at http://www.springer.com/series/10087

Gordana Dodig-Crnkovic
Raffaela Giovagnoli
Editors

Representation and Reality in Humans, Other Living Organisms and Intelligent Machines

Springer

Editors
Gordana Dodig-Crnkovic
Gothenburg
Sweden

Raffaela Giovagnoli
Faculty of Philosophy
Pontifical Lateran University
Vatican City
Italy

ISSN 2192-6255 ISSN 2192-6263 (electronic)
Studies in Applied Philosophy, Epistemology and Rational Ethics
ISBN 978-3-319-43782-8 ISBN 978-3-319-43784-2 (eBook)
DOI 10.1007/978-3-319-43784-2

Library of Congress Control Number: 2017943215

© Springer International Publishing AG 2017
This book was advertised with a copyright holder name “Springer International Publishing Switzerland” in error, whereas the correct copyright holder name is “Springer International Publishing AG”.
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This book project is based on several years of collaboration between two editors, starting with the organization of the symposium *Computing Nature* at the AISB/IACAP World Congress in 2012 in Birmingham. It continued with the AISB50 Convention at Goldsmiths, London in 2014 and the symposium *Representation: Humans, Animals and Machines* organized in cooperation with Veronica Arriola-Rios, University of Birmingham. During preparation, the book project was presented at the symposium *Representation and Reality* at UNILOG 2015, World Congress on Universal Logic in Istanbul. All those events, as well as our networks with research communities of cognitive scientists, computer scientists, philosophers, logicians, AI researchers, roboticists and natural scientists, connect us with the authors of the present volume, offering views on the topic of representation and its connections to reality in humans, other living organisms and machines. The choice of the title was made so as to refer to historical attempts at making connections, with Wiener's "*Cybernetics: Or Control and Communication in the Animal and the Machine*" (Wiener 1948), and Putnam's human-centric "*Representation and Reality*" (Putnam 1988).

How to address this vast topic and connect representation and reality in humans, other living beings and machines, based on the best contemporary knowledge? We invited prominent researchers with different perspectives and deep insights into the various facets of the relationship between reality and representation in those three classes of agents. How can we find a common link between reality-constructing agents like us humans with language ability and social structures that define our agency, other living organisms, from bacteria to plants and animals communicating and processing information in a variety of ways, with machines, physical and virtual? We have taken a cognitive, computational, natural sciences, philosophical, logical and machine perspective. Of course, no perspective is simple and pure, but rather a fractal structure in which recurrent mirroring of other perspectives at different scales and in different senses occurs. So in a contribution characterized predominantly as "cognitive perspective" there are elements of natural sciences, logic, philosophy and so on. Our aim is to provide a multifaceted view of the topic of representation and reality in the range of approaches from disciplinary (where a “discipline” is represented with its forefront outlooks) to multidisciplinary, interdisciplinary and cross disciplinary approaches. As a whole, the book presents a complex picture of a network of connections between different research fields addressing the state of the art of the topic reality versus representation.

We can see this book both as a collection of contributions of our authors to their fields of specialization and as an invitation to the reader to reflect on the kaleidoscope of domain-specific insights and their mutual connections in the context of the whole book. It is our common contribution to the continuous learning and shared knowledge about the nature of representation as found in the process of cognition (with “mind” as its philosophical reflection), cognition as it exists in different degrees in all living beings, including humans, and as it is recently being developed and constructed in machines in the field of cognitive computing.

Aspects of the relationship between representation and reality today connect networks of communities of philosophers, computer scientists, logicians, anthropologists, psychologists, sociologists, neuroscientists, linguists, information and communication scientists, system theoreticians and engineers, theoreticians and practitioners in computability and computing, information theory researchers, cybernetic systems researchers, synthetic biologists, biolinguists, bioinformaticians and biosemioticians, and many more. Current knowledge is distributed and no one’s insight is total and exhaustive, even though some can see more and broader while others see sharper and in more detail within specific perspectives. Here is a short presentation of different perspectives as represented in the book’s chapters, with an attempt to connect them into a common network.

Cognitive Perspectives

Connecting representation with reality, Terrence Deacon investigates the relation between information and reference, i.e. the relation of “aboutness”. He starts by making the distinction between the Shannon model of information and semantic information. The Shannon model describes communication of information and its goal is to engineer the best ways of communicating information through a noisy channel, based on the quantitative measure of the amount of information. However, in everyday use, the most important characteristic of information is its *meaning*, that is information about something relates to something in the world. Deacon’s analysis focuses on the capacity of a medium to provide reference and argues that this capacity can be seen in the difference between informational and thermodynamic entropy. This qualitative analysis shows that reference is a causally relevant physical phenomenon.

If representation is causally connected to reality, then how about hallucinations? In his contribution, Marcin Miłkowski studies models of visual hallucination in people with Charles Bonnet Syndrome and proposes to see them as illustrative cases where representations are not about anything in the real world. Miłkowski presents the computational model implementing neural network architecture with deep learning that can illuminate representational mechanisms for hallucination. It is interesting to observe that neural networks are often taken to be nonrepresentational models, in spite of the human brain being based on (biological) neural networks.

The old debate is still alive between representationalism and antirepresentationalism with the question of whether cognition relies on *representations mirroring* reality, or it is an *adaptive form of dynamics* based on the *interaction of an agent with the environment*. Typically, it is taken to be important for the critique of computational models of cognition as it is assumed that computations essentially depend on internal representations and that there is no computation without abstract symbol manipulation. This debate continues to this day because we still have no method to directly detect representations in a cognitive system.

In the opposition between *old (“orthodox”) computationalism* (based on the idea of the abstract logical universal Turing Machine) and *enactivism* (that emphasizes the central importance of the body and the environment for cognition), Tom Froese connects *meaning* for a living being with its embodiment with the (ever present) possibility of death. The claim this article makes is that *orthodox* computational models cannot account for meaning. Meaning as based in the necessity for active autopoiesis and the struggle for survival is outside the scope of the abstract Turing Machine model of computation.

Computationalism as a theory of mind is often criticized in its classical/orthodox approach, such as the computational representational theory of mind presented by Fodor. Jesus Ezquerro and Mauricio Iza propose an effective alternative *computational model of embodied cognition*, understanding language as predicting sensorimotor and affective effects of the action described by a verb.

Computational Perspectives

Dynamical systems are abstract mathematical objects that are used to describe physical processes. They are defined as “A means of describing how one state develops into another state over the course of time. Technically, a dynamical system is a smooth action of the reals or the integers on another object (usually a manifold)” by Wolfram MathWorld. Dynamical systems are frequently believed to be opposite to and irreconcilable with computational models. However, in their chapter Jan van Leeuwen and Jiří Wiedermann use exactly the dynamic systems formalism to develop a new *dynamic knowledge-based theory of computation* capable of explaining computational phenomena in both living and artificial systems.

Even though many people, even among scientists and philosophers, would say that the process of life and computation have nothing to do with each other, Dominic Horsman, Viv Kendon, Susan Stepney and Peter Young have dedicated their study to computation in living systems. Within the framework of Abstraction/Representation theory (AR theory), computing is assumed to be *representational* activity. In this chapter, the AR approach is used to elicit conditions under which a biological system computes. The framework is developed in the context of a nonstandard human-designed computing and has already been applied.

Nicolas Gauvrit, Hector Zenil and Jesper Tegner take a computational stance and cover in their chapter the whole range of the information-theoretic and algorithmic approaches to human, animal and artificial cognition. They review existing models of computation and propose algorithmic information-theoretic measures of cognition.

Unlike the above computational approaches to cognition, the article by Dean Petters, John Hummel, Martin Jüttner, Ellie Wakui, and Jules Davidoff is of more empirical character. Computational models of object recognition are used to investigate representational change in the course of development in humans. By means of developmental studies in children, and computational modeling of their results by artificial neural networks, in comparison with the existing research on adults, it was possible to follow how visual representations mediate object recognition.

Natural Sciences Perspectives

Reductionism is one of the most severe sins one can commit, according to many philosophers and cognitive scientists. One should not even try to investigate a physical substrate on which cognition definitely relies. Gianfranco Basti is obviously untroubled by the danger of being accused of reductionism, and he examines the deepest roots of information processing in the physical world—the quantum field theory as a dual paradigm in fundamental physics—connecting it with semantic information in cognitive sciences. Basti makes us aware of the paradigm shift with respect to the Standard Model of physics, as well as quantum mechanics conceived as the many-body-dynamics generalization from classical mechanics. The basic assumption of the old paradigm about the closed physical system does not hold, as quantum field theory systems are inherently open to the background fluctuations of the quantum vacuum, and are capable of system phase transitions. It is interesting to observe that Prigogine had that insight in thermodynamics, making the step from closed thermodynamical systems or systems in thermodynamical equilibrium to the study of open thermodynamical systems with the inflow of energy that exhibited stable self-organized patterns. Maturana and Varela's work further developed our understanding of autopoiesis in living cells as processes essentially dependent on the openness of the system to the exchanges and interactions with the environment.

Irrespective of the representationalism versus antirepresentationalism discussion, today we have a broader concept of computation (computing nature) which posits that *every natural system computes*, with computation as its physical dynamics as presented in the chapter by Gordana Dodig-Crnkovic and Rickard von Haugwitz. That would mean that one does not need to search for representations in the brain just for the sake of settling the debate on whether the brain computes. Being a natural system, the human brain computes as well as the rest of nature in the framework of computing nature.

Dodig-Crnkovic and von Haugwitz search for answers to the questions: What is reality for an agent? What is minimal cognition? How does the morphology of a cognitive agent affect cognition? How do infocomputational structures evolve from the simplest living beings to the most complex ones? As a framework for answering these questions an infocomputational nature is assumed, constructed as a synthesis of (pan)informational ansatz (structures in nature are informational structures for us as cognitive agents) and (pan)computational view (dynamics of nature is computation), where information is defined as a structure (for an agent), and computation as the dynamics of information (information processing). Both information and computation in this context have broader meaning than in everyday use, and both are necessarily grounded in physical implementation, and the aim of this approach is to integrate computational approaches with embodiment and enactivism.

Philosophical Perspectives

Raffaela Giovagnoli presents the theme of language, classically central for the problem of representation. She focuses on inferential linguistic practices, which characterize human cognition. A pragmatic account helps us to understand what representational capacities are peculiar to humans, animals and machines. A study of the history of concept formation and use enables a clear distinction between humans and other animals. Starting from Frege, Giovagnoli presents the notion of representation theorized by Searle, and introduces the ideas proposed by the Brandom pragmatist order of explanation. The author points to the clear distinction between human and animal linguistic practices that can also be applied to the study of linguistic practices in machines.

Angela Ales Bello deals with the problem of consciousness in humans, animals and machines in a phenomenological account based on the hyletic dimension. The concept of hylé (a transliteration of Aristotle's concept of *matter*) was proposed by Husserl to denote the *nonintentional*, *direct sensory* aspects of living experience (*qualia*), shared by humans and animals. Ales Bello wishes to launch a challenge to those positions that ground themselves in the presumed objectivity of modern science. Starting from Husserl, Ales Bello examines the essential elements of phenomenology in order to understand how they developed and how they stand in relation to various scientific views. In this context the question "Can we conclude that the human being is like a machine?", looking at machines of today, is answered in the negative.

One of the arguments against computational models of mind is the so-called "hard problem of consciousness", which argues for irreducibility of the qualitative phenomenal features of mental experience to its functional cognitive features. Roberta Lanfredini presents the argument that "not only the classical cognitive pattern but also the classical phenomenological pattern give rise to a problem concerning the qualitative dimension." There is not only "the hard problem of consciousness" but "the hard problem of matter" as well. Lanfredini connects *representation* with *matter* and emphasizes the *central importance of time*, and *dynamics* in the phenomenology of the mind.

Logical Perspectives

The question of capturing homogeneity and heterogeneity is central for representation. Bateson's definition of information as a difference that makes a difference (for an agent) can be applied not only to perception but also to reasoning. How do we define differences and their role in reasoning? Henri Prade and Gilles Richard address comparative thinking as a basis of our comprehension of reality and present a description of logical proportions in comparing two objects or situations in terms of Boolean features within a cube of opposition. Homogeneous and heterogeneous logical proportions are important for classification, completion of missing information and anomaly detection, and they provide insights into human reasoning.

As already mentioned, among the strongest and recurrent objections against computational models of cognition is their alleged incapability to account for meaning (semantics). Ferdinando Cavaliere offers a remedy for this problem in the case of online search engines by means of adding the semantics with the help of a hypothetical program "Semantic Prompter Engine", based on the "Distinctive Predicate Calculus" logic, designed by the author. The aim of this approach is to bridge the gap between natural and artificial representations of concepts and reality by providing the semantics for the latter.

In his chapter Jean-Yves Béziau continues this tradition of critical analysis of human rationality, offering an original comparison between human and animal intelligence. Rationality in language is the logical way we have to represent knowledge but the richness of our emotional life is worth further exploration.

Machine Perspectives

Even Matej Hoffmann and Vincent Müller, like Jan van Leeuwen and Jiří Wiedermann in their chapter, find the dynamical systems framework as the most suitable, in their case for connecting computation controllers and robotic bodies. They introduce the concept of *morphological computation* in the sense it is used in robotics, which refers to "offloading" computational processing from a central controller to the body of a robot, exploiting the use of self-organization of the physical system in its interaction with the environment.

David Zarebski distinguishes between three types of ontology: the mind-independent structure of reality studied in philosophy and formal ontology, the structure of the human representation of the world studied in the cognitive sciences, or the structure of knowledge representation investigated in data engineering. Zarebski describes interactions between those three research fields – philosophy with formal ontology, cognitive science and data and knowledge engineering – as consisting in cognitive science explaining how human cognitive capacities can affect metaphysics, or in what way information systems ontologies can learn from formal ontologies. While cognitive science typically makes no distinction between ontological and epistemological realism, Zarebski defends the position that it is possible, in the Kantian tradition, to be realistic about the epistemology without the need for naturalization of the ontology.

A basic question regarding intelligent machinery concerns the nature of machine intelligence. Philip Larrey's chapter compares human intelligence and machine intelligence, as it looks today and as anticipated in the future. Based on Bostrom's classification of future-envisaged machine superintelligence into speed-, collective- and quality-superintelligence, Larrey presents critical views of all three types, proposing a fourth type that would combine human and machine intelligence, under the assumption that humans will do the really intelligent part while machines will continue as before to provide different services without being conscious of what they are doing. It is hard not to agree with the author's conclusion: "Ours is truly an 'unknown future'".

It remains to see what future developments of machine intelligence, robotics and cognitive computing will bring, and if perhaps in the future machines get one more capacity, "machine consciousness", which will not be like human consciousness, in the same way that "machine learning" is not like human learning, and machine walking is not like human walking but fulfills that function for a machine. Larrey's chapter considers the possibility of building (super)intelligent conscious machines. The next question, whether it is a good idea, or under what constraints is it justified to build possibly (super)intelligent conscious artifacts, is a different one, addressed both by Bostrom as well as Hawkins, Tegmark, Boden and many other prominent scientists. It might be considered a topic for ethicists or for political decision-makers, but the existing knowledge of those communities must be constantly updated through insights from researchers dealing in depth with the phenomena of natural and artifactual cognitive systems.

To sum up, the aim of this book is to enrich our views on representation and deepen our understanding of its different aspects. It is seldom the case that one discipline can exhaust all the complexity of a real-world phenomenon. The historical divisions formed deep trails that even the coming generations of researchers tend to follow, and existing divisions built into academic institutions, publications and research funding often present obstacles that prevent us from researching the big complex picture and our possible role in the overall network of knowledge. We are trying to provide a glimpse behind the mirror of our own specialist views on the phenomenon of representation and its connections to reality, written by researchers with different commitments and preferences regarding divisions into the computationalist versus enactivist approach, cognitive science versus phenomenology, logic versus. emotions, the "hard problem of mind" versus the "hard problem of matter”, and so on. It should be emphasized that in spite of seemingly impermeable barriers between well-known opposing choices there is a natural building of networks that produces hybrids—enactivist studies based on computational simulations and physical computation models that take embodiment as fundamental, computational models of dynamical systems as well as dynamic systems formulations of computational frameworks, reductionism with the character of deconstruction, opening new views on emergence, the list goes on. We hope with this to contribute to the dialogue and further mutual connections between the research communities involved.

In conclusion we want to thank our authors for their excellent contributions, as well as for their involvement in the open and completely transparent review process done in a collegial and constructive spirit, where each chapter got at minimum three and at most eight well-informed and helpful reviews. We are thankful to Robert Lowe for his contribution to the review process.

Last but not least we want to thank our publisher, Ronan Nugent at Springer, for his continuous support and friendly advice in this project.

Gothenburg, Sweden Gordana Dodig-Crnkovic
Vatican City, Italy Raffaela Giovagnoli

Contents

Part I Cognitive Perspectives

Part II Computational Perspectives

Part III Natural Sciences Perspectives

Part IV Philosophical Perspectives

Part V Logical Perspectives

Part VI Machine Perspectives

Contributors

Gianfranco Basti Faculty of Philosophy, Pontifical Lateran University, Rome, Italy

Angela Ales Bello Pontifical Lateran University Rome, Rome, Italy

Jean-Yves Beziau University of Brazil, Rio de Janeiro, Brazil; University of San Diego, San Diego, CA, USA

Ferdinando Cavaliere Circolo Matematico Cesenate, Cesena, (FC), Italy

Jules Davidoff Goldsmiths College, London, UK

Terrence W. Deacon University of California, Berkeley, USA

Gordana Dodig-Crnkovic Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

J. Ezquerro Institute for Logic Cognition Language and Information (LCLI), University of the Basque Country, UPV/EHU, Donostia-San Sebastián, Spain

Tom Froese Institute for Applied Mathematics and Systems Research, National Autonomous University of Mexico, Mexico City, Mexico; Center for the Sciences of Complexity, National Autonomous University of Mexico, Mexico City, Mexico

Nicolas Gauvrit Human and Artificial Cognition Lab, Université Paris 8 & EPHE, Paris, France

Raffaela Giovagnoli Faculty of Philosophy, Lateran University, Rome, Italy

Matej Hoffmann Faculty of Electrical Engineering, Department of Cybernetics, Czech Technical University in Prague, Prague, Czech Republic; iCub Facility, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genoa, Italy

Rickard von Haugwitz Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden

Dominic Horsman Department of Computer Science, University of Oxford, Oxford, UK ; Department of Physics, Durham University, Durham, UK

John Hummel University of Illinios at Urbana-Champaign, Champaign, USA

M. Iza Department of Psychology, University of Malaga, Málaga, Spain

Martin Jüttner University of Aston, Birmingham, UK

Viv Kendon Department of Physics, Durham University, Durham, UK

Roberta Lanfredini Dipartimento di Lettere e Filosofia, Florence, Italy

Philip Larrey Pontifical Lateran University, Vatican City, Italy

Marcin Miłkowski Institute of Philosophy and Sociology, Polish Academy of Sciences, Warszawa, Poland

Vincent C. Müller Anatolia College/ACT, Pylaia, Greece; Department of Philosophy, IDEA Centre, University of Leeds, Leeds, UK

Dean Petters Birmingham City University, Birmingham, UK

Henri Prade IRIT, Université Paul Sabatier, Toulouse Cedex 09, France

Gilles Richard IRIT, Université Paul Sabatier, Toulouse Cedex 09, France

Susan Stepney Department of Computer Science and York Centre for Complex Systems Analysis, University of York, York, UK

Jesper Tegnér Unit of Computational Medicine, Department of Medicine Solna, Centre for Molecular Medicine & SciLifeLab, Karolinska Institutet, Stockholm, Sweden

Jan van Leeuwen Department of Information and Computing Sciences, Utrecht University, CC Utrecht, The Netherlands

Elley Wakui University of East London, London, UK

Jiří Wiedermann Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague 8, Czech Republic; Czech Institute of Informatics, Robotics, and Cybernetics, Czech Technical University in Prague, Prague 6, Czech Republic

J.P.W. Young Department of Biology, University of York, York, UK

David Zarebski IHPST, University Paris 1 Pantheon Sorbonne, Paris, France

Hector Zenil Unit of Computational Medicine, Department of Medicine Solna, Centre for Molecular Medicine & SciLifeLab, Karolinska Institutet, Stockholm, Sweden; Department of Computer Science, University of Oxford, Oxford, UK

Part I

Cognitive Perspectives

Information and Reference

Terrence W. Deacon

Abstract The technical concept of information developed after Shannon [22] has fueled advances in many fields, but its quantitative precision and its breadth of application have come at a cost. Its formal abstraction from issues of reference and significance has reduced its usefulness in fields such as biology, cognitive neuroscience and the social sciences where such issues are most relevant. I argue that explaining these nonintrinsic properties requires focusing on the physical properties of the information medium with respect to those of its physical context—and specifically the relationship between the thermodynamic and information entropies of each. Reference is shown to be a function of the thermodynamic openness of the information medium. Interactions between an informing medium and its physical context that drive the medium to a less probable state create intrinsic constraints that indirectly reflect the form of this extrinsic influence. This susceptibility of an informing medium to the effects of physical work is also relevant for assessing the significance or usefulness of information. Significance can be measured in terms of work "saved" due to access to information about certain contextual factors relevant to achieving a preferred target condition.

1 Introduction

> *I didn't like the term Information Theory. Claude didn't like it either. You see, the term 'information theory' suggests that it is a theory about information – but it's not. It's the transmission of information, not information. Lots of people just didn't understand this… information is always about something. It is information provided by something, about something. (Interview with R. Fano, 2001)*

> *What I have tried to do is to turn information theory upside down to make what the engineers call "redundancy" [coding syntax] but I call "pattern" into the primary phenomenon… (Gregory Bateson, letter to John Lilly on his dolphin research, 10/05/1968)*

T.W. Deacon (✉)
University of California, Berkeley, USA
e-mail: deacon@berkeley.edu

G. Dodig-Crnkovic and R. Giovagnoli (eds.), *Representation and Reality in Humans, Other Living Organisms and Intelligent Machines*, Studies in Applied Philosophy, Epistemology and Rational Ethics 28, DOI 10.1007/978-3-319-43784-2_1

In common use and in its etymology the term "information" has always been associated with concepts of reference and significance—that is to say it is *about* something for some *use*. But following the landmark paper by Claude Shannon [22] (and later developments by Wiener, Kolmogorov and others), the technical use of the term became almost entirely restricted to refer to signal properties of a communication medium irrespective of reference or use. In the introduction to this seminal report, Shannon points out that, although communications often have meaning, "These semantic aspects of communication are irrelevant to the engineering problem," which is to provide a precise engineering tool to assess the computational and physical demands of the transmission, storage and encryption of communications in all forms.

The theory provided a way to precisely measure these properties as well as to determine limits on compression, encryption and error correction. By a sort of metonymic shorthand this quantity (measured in bits) came to be considered synonymous with the meaning of "information" (both in the technical literature and in colloquial use in the IT world) but at the cost of inconsistency with its most distinctive defining attributes.

This definition was, however, consistent with a tacit metaphysical principle assumed in the contemporary natural sciences: the assertion that only material and energetic properties can be assigned causal power and that appeals to teleological explanations are illegitimate. This methodological framework recognizes that teleological explanations merely assign a locus of cause but fail to provide any mechanism, and so they effectively mark a point where explanation ceases. But this stance does not also entail a denial of the reality of teleological forms of causality nor does it require that they can be entirely reduced to intrinsic material and energetic properties.

Reference and significance are both implicitly teleological concepts in the sense that they require an interpretive context (i.e. a point of view) and are not intrinsic to any specific physical substrate (e.g. in the way that mass and charge are). By abstracting the technical definition of information away from these extrinsic properties, Shannon provided a concept of information that could be used to measure a formal property that is inherent in all physical phenomena: their organization. Because of its minimalism, this conception of information became a precise and widely applicable analytic tool that has fueled advances in many fields, from fundamental physics to genetics to computation. But this strength has also undermined its usefulness in fields distinguished by the need to explain the non-intrinsic properties associated with information. This has limited its value for organismal biology, where function is fundamental, for the cognitive sciences, where representation is a central issue, and for the social sciences, where normative assessment seems unavoidable. So this technical redefinition of information has been both a virtue and a limitation.

The central goal of this essay is to demonstrate that the previously set aside (and presumed nonphysical) properties of reference and significance (i.e. normativity) can be reincorporated into a rigorous formal analysis of information that is suitable for use in both the physical (e.g. quantum theory, cosmology, computation theory) and semiotic sciences (e.g. biology, cognitive science, economics). This analysis builds on Shannon's formalization of information but extends it to explicitly model its link to the statistical and thermodynamic properties of its physical context and to the physical work of interpreting it. It is argued that an accurate analysis of the nonintrinsic attributes that distinguish information from mere physical differences is not only feasible, but necessary to account for its distinctive form of causal efficacy.

Initial qualitative and conceptual steps toward this augmentation of information theory have been outlined in a number of my recent works [8–12]. In these studies we hypothesize that both a determination of reference and a measure of significance or functional value can be formulated in terms of how the extrinsic physical modification of an information-bearing medium affects the dynamics of an interpreting system that exhibits intrinsically end-directed and self-preserving properties.

2 Background

The problems posed by the concepts of reference and significance have been the subject of considerable philosophical debate over many centuries. It would probably not be hyperbole to suggest that they are among the most subtle and complex issues in the history of philosophy. So it might seem presumptuous to imagine that these questions can be settled with a modest extension of formal information theory. Though it is often assumed that the current technical treatment of information either resolves these issues or shows them to be irrelevant, the fact that concepts of meaning and interpretation are treated as irrelevant to physical science and in biology indicates that a major gap in understanding still separates these technical uses from the traditional meaning of the term. Though it is not possible to even come close to doing justice to this long and byzantine philosophical history of analysis in this essay, it is worth setting the stage with a fundamental insight gleaned from this perspective.

The challenge posed by these ideas was eloquently, if enigmatically, formulated by Franz Brentano's [4] use of the terms "intention" and "inexistence" (borrowed from Medieval philosophy) when describing the referential property of information.

> Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity

… Brentano's use of the curious term "inexistence" points to the fact that informational content is not exactly an intrinsic physical property of the medium that conveys it, even though it somehow inheres in it as well. The reference conveyed is typically displaced, abstract, ambiguous, or possibly about something nonexistent. And in any case it is dependent on interpretation. Excluding this attribute from the technical concept of information appears to have been necessary to ground information theory in physics and make it suitable for engineering applications, but *this was bought at a cost of ignoring the very attributes that distinguish information from other physical phenomena*.

I argue that the key to formulating a more adequate concept of information that includes these "inexistent" properties is to be found, ironically, in more carefully attending to the physicality of information media. A hint that this is important is captured in two distinctively different uses of the concept of entropy (informational entropy and thermodynamic entropy). The term "entropy" was, of course, originally coined by Rudolph Clausius in [7] in the context of the early development of a concept of the relationship between heat and physical work. In 1874 Ludwig Boltzmann [3] formalized it further in his H-theorem that represents entropy in terms of the sum of the probabilities of microstates of a dynamical system. In Shannon's [22] analysis (building on the work of his predecessors at Bell Labs, Nyquist and Hartley), this same formula is used to represent the information of a given communication medium, also describing its distinguishable states. Apparently, following a suggestion from the mathematician John von Neumann, he also called this value "entropy". Many writers have cogently argued that these two concepts should not be confused (e.g. [15, 20, 24, 26]), and that Shannon's choice of this term to describe a statistical property of information media was a colossal mistake (e.g. [25]). And there is, of course, much to distinguish these two uses of the same term beyond this abstract mathematical similarity. For example, there is no informational analogue to Clausius' theorem that the total entropy of an isolated physical system can only increase (the second law of thermodynamics). Nor is informational entropy a dynamical concept associated with energy and work. Nevertheless, many have attempted to discern a deeper linkage underlying these statistical analogies (e.g. [2, 14, 21]).

The idea that there is a link between information and thermodynamics has a long history. Ways of demonstrating such a linkage have been proposed in many forms from Maxwell's [19] famous demon [18] to Landauer's [17] argument about the thermodynamic cost of information erasure. However, these approaches focus primarily on the energetic and thermodynamic "costs" of manipulating physical markers or taking measurements and how this might alter system entropy. So they have largely ignored the problem of how reference and significance are physically instantiated, to instead defend the belief that information cannot be used to violate the second law of thermodynamics.

Both notions of entropy can be applied to physical features of an information medium. This is because any physical medium capable of conveying information must be able to exhibit different states. As Shannon demonstrated, the number and relative probabilities of these different states (and, in a continuous communication, their rates of change) are what determines the capacity for that medium to "store" or "convey" information. Following insights from statistical mechanics, he argued that this could be measured in terms of the value of $-\sum p_i \log p_i$, where p_i is the probability of the occurrence of a given state i of the information-bearing medium. Any process, structure, or system that can be analyzed onto component states, each able to be assigned a probability of being exhibited, can in this way be described in terms of this measure of information. This insight at the foundation of Shannon's approach also provided one of the first widely accepted model-independent measures of complexity [11]. Most subsequent means of measuring form or complexity have been developed with respect to this basic insight (as in the work of [6, 16], etc.). Information theory thus became more than just a tool for analyzing communication.

Besides providing a measure of the information capacity of a given medium, Shannon's analysis also demonstrated that the amount of information provided by a *received* signal (i.e. a message) can be measured as a function of the amount of uncertainty that the received signal removes. This can be measured as a difference between the prior (or intrinsic) uncertainty (the Shannon entropy) of the signal medium being in a given state and its current received state (e.g. in a received message). This necessarily relational nature of information is an important distinction that is often overlooked, and it effectively distinguishes two interdependent and inter-defined uses of the concept: the potential information capacity of a given medium and the information provided by a specific message conveyed by that medium. Both states can be assigned an entropy value. Using this relative measure, problems of noise, error correction and encryption can be likewise analyzed in terms of differences or changes in component signal state probabilities.

The relational nature of information, even in this technically minimalistic form, is an important clue to the fact that information is not a simple intrinsic property of things. Interestingly, this is also a feature that both concepts of entropy share. The third law of thermodynamics likewise asserts that entropy is a relational measure: a difference between states of a system. The relational nature of thermodynamic entropy was not fully appreciated until 1906 when Walter Nernst augmented thermodynamic theory establishing an absolute reference point at 0° K. Likewise the relational nature of informational entropy is also often overlooked.

The entropy of a signal is also not an intrinsic property but a relational property. This relational character of both concepts of entropy is a clue that the statistical signal properties of a medium are linked to its ability to convey reference. Both Shannon's analysis and thermodynamic theory depend on comparing the relative degree of constraint on current entropy with respect to what is minimally and/or maximally possible for a given system. A received signal that exhibits constraint in its information entropy is more predictable even if it is not reduced to a single fixed value. Thus even a noisy signal reduces uncertainty by this difference, so long as it is not fully random.

Noise, which is the corruption of a signal, increases uncertainty by reducing this constraint. But noise is often the result of physical degradation of the conveying medium. This is yet a further clue to the intrinsic interrelationship between thermodynamic and information entropy.

The distinction between signal and noise is an important clue to how an information-bearing medium can be linked to some nonintrinsic object, event, or property. Noise is a difference or change of the information entropy of a message. Often, as in the case of radio transmission, noise is due to an increase in thermodynamic entropy of the conveying medium as a result of interference or simple signal degradation, i.e. to factors extrinsic to the information-bearing medium that affect its physical attributes. Any such physical influence will entail physical work, and work entails a change in thermodynamic entropy. This suggests that what at first appears to be an unrelated and superficial parallelism between these two concepts of entropy is instead a critical clue to how a signal medium can be about something that it is not.

Noise and signal are *both* linked to something extrinsic that has affected the information medium. The difference is due to interpretive assessment, not anything intrinsic to the informing medium. The nonintrinsic distinction between signal and noise is an additional relational attribute indicating that, even though specific reference and significance can be set aside to analyze the statistical properties of an information-bearing signal, they are nevertheless assumed.

3 Physicality of Aboutness

To exemplify the way that the statistical properties of a medium can provide the potential for reference, consider the use of information-theoretic analyses in molecular biology. The statistical properties of nucleotide sequences can provide critical clues to potential biological functions even though the sequencing of genomes is largely accomplished in ignorance of any specific gene function. Indeed, exploring the statistical structure of these sequences has provided many important clues to unanticipated functional properties of DNA, RNA and the organic properties they specify. For example, the presence of constraints on the possible statistical entropy of a gene sequence, exhibited in sequence redundancy across species, tends to predict that the sequence in question plays a functional role in the organism, i.e. that it encodes information “about” that function. This is because constraint in the form of redundancy provides evidence that nonrandom—i.e. functional—influences are at work. Simply scanning gene sequences for constraint, then, provides a tool for discovering sequences that probably contain information “about” some function, even though no specific functional information is provided.

The concept of mutual information is also a clue to the way that signal constraint relates to reference. The mutual information between two signals or sequences of alphanumeric characters is a measure of their statistical nondifference. This can even be assessed despite nonidentity of any of the components in the two, as for example exists in the statistical parallels between nucleotide sequences in DNA and amino acid sequences in proteins. Shannon demonstrated that any degree of signal noise less than total noise can be compensated for by a comparable level of signal redundancy. In this respect, redundancy—a constraint on possible variety—is what preserves signal reference. In genetic evolution, sequence redundancy and mutual information are clues that a given genetic signal carries functional information and provides a kind of reference to cellular-molecular dynamics that were useful in the past and are likely useful in the present. As the epigraph to this essay from Gregory

Bateson suggests, reference is ultimately embodied in signal constraint (i.e. redundancy and the regularity or pattern that this produces).

So implicit in the technical sense of "information" is an understanding that statistical properties of the signal medium are in some way related to its referential function. The question that I want to address is how it might be possible to more precisely characterize this relationship.

To address this let us begin with a focus on the relationship between thermodynamic and information entropy. Worries about the relationship between information and thermodynamics emerged in parallel with the kinetic theory of gasses and long predate Shannon's use of the entropy concept. They are best exemplified by the many analyses that developed in response to Maxwell's [19] famous thought experiment, which subsequently came to be known as Maxwell's demon. Though he was one of the major contributors to the formalization of the second law of thermodynamics, Maxwell pondered the possibility that molecular-level information might be able to drive an otherwise isolated thermodynamic system from a higher to a lower entropy state in violation of the second law. He described this problem in terms of a fanciful microscopic "demon" able to judge the relative momentum and/or velocity of gas molecules in either side of a two-chambered container with a closable passage separating them. He wondered whether the demon could use these measurements to determine whether or not to let a molecule move from one container to the other. Such a demon (or a device that acted in the same way) could selectively allow only fast-moving molecules to pass one way and only slow-moving molecules to pass the other way, thus progressively reducing the entropy of the whole system, in violation of the second law. But if this were even logically conceivable (even if not achievable with any physical mechanism), then the ubiquity and ineluctable directionality of the second law would be questioned.

Subsequent analysts have variously attempted to answer this conceptual challenge and preserve the second law by analyzing the thermodynamic aspects of the information assessment process. This has been explored in simplified abstract mechanical terms involving one or just a few moving particles (e.g. [23]), or by considering the relative entropic "cost" of acquiring this information (e.g. [5]), or in terms of the need to erase information from previous measurements in order to repeat the process (e.g. [17]). The intention has been to show that, even if one were able to create a device capable of such actions, the second law of thermodynamics would be preserved because there would be more entropy generated by obtaining and using this information than would be reduced by its actions.

In contrast, I propose to analyze the referential property of information directly and irrespective of the energetic cost of generating or erasing bits of data (though ultimately this cost is a relevant factor as well). For this purpose I instead focus on one of the most basic principles of Shannon's analysis: the role played by statistical constraint in the assessment of the amount of uncertainty removed by receipt of a given message or introduced by thermodynamic noise in a transmission. This allows us to make an abstract, but direct, comparison between the entropy of a given information-bearing medium in Shannon's terms and the entropy of that same medium in physical terms (though not necessarily just in thermodynamic terms).

The unifying factor is that both are expressions of the physical attributes of the information-bearing medium.

The essential point is this: every medium for storing or conveying information is constituted physically and its distinguishable states are physical states. So any change in that medium's statistical physical properties (e.g. its thermodynamic entropy) could potentially also change its informational properties. This is not a necessary relationship, since the distinguishable states used to convey information in any given case are inevitably a very small subset of the total range of different states that the physical medium can assume. However, because of its physicality, any change in the informational entropy of a given medium must necessarily also entail a change in its physical statistical properties. And, following the strictures of the second law of thermodynamics, any physical medium will only tend to be in an improbable constrained state if it has in some way been driven away from its more probable state by the imposition of physical work or prevented from achieving it by some extrinsic restriction. In other words, the relationship between the most probable state of the medium and the observed state at any particular moment is a reflection of its relationship with its physical context. Its intrinsic statistical properties are therefore clues to factors that are extrinsic to it.

In this analysis I build on this insight to argue that *referential* information is based on the constraints generated by physical work introduced due to the thermodynamic openness of an information-bearing medium and its susceptibility to contextual modification. The next section provides a point-by-point outline of the logic that formally defines the referential aspect of information in terms of the relationship between Shannon (information) entropy and Boltzmann (thermodynamic) entropy.

4 Steps to a Formalization of Reference

A. General case: passive information medium near equilibrium (e.g. geological formation, crime scene evidence, data from a scientific experiment, text, etc.)

1. Information (e.g. Shannon) entropy is not equivalent to thermodynamic (e.g. Boltzmann–Gibbs) entropy (or to the absolute statistical variety of physical states). [For convenience these entropies will be provisionally distinguished as Shannon vs. Boltzmann entropy, though recognizing that each includes multiple variant forms.]
2. However, for any physical signal medium, a change in Shannon entropy must also correspond to a change in Boltzmann entropy, though not vice versa because the distinctions selected/discerned to constitute the Shannon entropy of a given signal medium are typically a small subset of the possible physical variety of states—e.g. statistical entropy—of that medium.

3a. The Shannon information of a received message is measured as a reduction of signal uncertainty (= a reduction of Shannon entropy).

3b. For a simple physical medium, reduction of Shannon entropy must also correspond to a reduction of the Boltzmann entropy of that medium.
3c. This can be generalized as "any deviation away from a more probable state" (which can violate 3b in the case of media that are actively maintained in an improbable state, such as maintained far from equilibrium). (**See B below**.)
4a. A reduction of Boltzmann entropy of any physical medium is exhibited as constraint on its possible states or dynamical "trajectories".
4b. The production of physical constraint requires physical work in order to produce a decrease of Boltzmann entropy, according to the second law of thermodynamics.
5a. For a *passive medium* the physical work required to reduce its Boltzmann entropy must originate from some physical source *extrinsic* to that medium.
5b. Generalization: Constraint of the Shannon entropy of a passive medium = constraint of its Boltzmann entropy = the imposition of prior work from an external source.

6. An increase in constraint (i.e. deviation away from a more probable state) in the information medium literally "re-presents" the physical relationship between the medium and the extrinsic contextual factors (work) that caused this change in entropy (= what the information embodied in the constraint can be "about").
7. Since a given constraint has statistical structure, its *form* is a consequence of the specific structure of the work that produced it, the physical susceptibilities of the information-bearing medium and the possible/probable physical interactions between that medium and this extrinsic contextual factor.
8. The form of this medium constraint therefore corresponds to and can indirectly "re-present" the *form* of this work (i.e. in-*form*-ation).
9. Conclusion 1. The possibility of reference in a passive medium is a direct reflection of the possibility of a change in the Boltzmann and Shannon entropies of that medium due to a physical interaction between the information-bearing medium and a condition extrinsic to it.
10. Conclusion 2. The possible range of contents thereby referred to is conveyed by the form of the constraint produced in the medium by virtue of the form of work imposed from an extrinsic physical interaction.
11. Conclusion 3. The informing power of a given medium is a direct correlate of its capacity to exhibit the effects of physical work with respect to some extrinsic factor.
12. Corollary 1. What might be described as the referential entropy of a given medium is a function of the possible independent dimensions of kinds of extrinsically induced physical modifications it can undergo (e.g. physical deformation, electromagnetic modification, etc.) multiplied by the possible "distinguishable" states within each of these dimensions.

13. Corollary 2. Having the potential to exhibit the effects of work with respect to some extrinsic physical factor means that even no change in medium entropy or being in a most probable state still can provide reference (e.g. the burglar alarm that has not been tripped, or the failure of an experimental intervention to make a difference). It is thus reference to the fact that *no work to change the signal medium has occurred.*

In addition, since not all information-bearing media are inert physical structures or otherwise passive systems at or near thermodynamic equilibrium, we need to modify certain of these claims to extend this analysis to media that are themselves dynamical systems maintained far from equilibrium. This yields the following additional claims:

B. Special case: nonpassive information medium maintained far from equilibrium (e.g. metal detector or organism sense organ)
1. A persistently far-from-equilibrium process is one that is maintained in a lowered probability state. So certain of the above principles will be reversed in these conditions, specifically, those that depend on extrinsic work moving a medium to a lower probability, lower entropy state.
2. Maintenance of a low Boltzmann entropy dynamical process necessarily requires persistent physical work or persistent constraints preventing an increase of Boltzmann and Shannon entropies.
3. Any corresponding increase in Shannon entropy therefore corresponds to a disruption of the work that is maintaining the medium in its lower entropy state. This can occur by impeding the intrinsic work or disrupting some dissipation-inhibiting constraint being maintained in that system.
4a. An increase in the Shannon entropy of a persistently far-from-equilibrium information medium can thereby “indicate” extrinsic interference with that work or constraint maintenance.
4b. A persistently far-from-equilibrium dynamical medium can be perturbed in a way that increases its entropy due to contact with a passive extrinsic factor. Any passive or dynamic influence that produces a loss of constraint in such a system can provide reference to that extrinsic factor.
5a. Since work requires specific constraints and specific energetic and material resources, these become dimensions with respect to which the change in entropy can refer to some external factor.
5b. The dynamical and physical properties of a far-from-equilibrium information-bearing medium determine its “referential entropy”.
6. Corollary 3. This can be generalized to also describe the referential capacity of any medium normally subject to regular end-directed influences that tend to cause it to be in an improbable or highly constrained state. This therefore is applicable to living systems with respect to their adaptations to avoid degradation and also to far more complex social and cultural contexts where there is active “work” to maintain certain “preferred” orders.

5 Active Acquisition of Information

The far-from-equilibrium case is of major importance also because it provides the foundation for an analysis of the nature of an interpretive process, such as might be applied to simple organism adaptations and genetics (see next section).

A simple exemplar of a far-from-equilibrium information medium is a metal detector. Metal detectors typically operate by constantly maintaining a stable electromagnetic field. This signal medium requires the work provided by the constant flow of an electric current through a coil. This magnetic field is easily distorted by the presence of a conducting object, thus interfering with the electronic work otherwise maintaining a redundant signal also linked to this work (often the tone produced by an oscillator sensitive to a change in current).

This is roughly analogous to the way that living processes gain information about their world. Whether by virtue of a ligand binding to a specific receptor molecule on a cell surface and changing its conformation or a photon modifying the dynamics of signal processing in retinal neurons, it is ultimately interference with an ongoing living process requiring metabolic work that provides referential information to an organism.

Recall that, according to Shannon's analysis, the measure of information in a message is proportional to the reduction of entropy in the received signal compared with the potential entropy of the channel (i.e. the medium). Now we can see that reduction of the Boltzmann entropy ($-\Delta S_b$) of the passive information medium is proportional to a reduction of its Shannon entropy ($-\Delta S_s$) by some proportionality constant m (usually far below 1.0) that determines what portion of its Boltzmann entropy is used as Shannon entropy. In the far-from-equilibrium case the situation is reversed: some fraction of the increase in a medium's Boltzmann entropy ($+\Delta S_b$) is proportional to an increase of its Shannon entropy ($+\Delta S_s$). In both cases, information about the source of an external disturbance is negatively embodied in the Shannon entropy ($\pm\Delta S_s$) of the medium as a change in intrinsic constraint.

In general this means that a medium in its most probable state, exhibiting no change in entropy, can also provide information about the *absence* of a given specific referent. This is nevertheless a form of reference. A smoke alarm that remains silent in the absence of smoke still provides information. So both a high probability state and a low probability state of an information medium are potentially referential, demonstrating that every referential relationship corresponds to a reduced probability state of Shannon entropy. Thus we can conclude that reference is made possible by the susceptibility of a given information medium to reflect the effect of work with respect to an extrinsic context, and that the sign of this effect—i.e. whether there is an increase or decrease of medium constraint—will depend on whether this work originates in the interpretive process or in its extrinsic physical context.

6 Conclusion

The above analysis demonstrates that the capacity of an informing medium to provide reference (i.e. "aboutness") derives from a linkage between its information entropy and thermodynamic entropy. Despite the fact that these are nonequivalent statistical measures of different properties (e.g. free energy vs. formal complexity, respectively) they can both reflect related changes introduced by physical work. So although reference is not an intrinsic physical property of any information-bearing medium, this linkage to work renders the referential property of information susceptible to exact formal and empirical analysis. Thus reference is not a subjective, heuristic, or epiphenomenal product of prescientific theorizing, like phlogiston, able to be dispensed with as more precise physical science comes to explain it away. Rather, it is a causally relevant physical property affecting systems that depend on selected contextual features for their operation, but lack direct access to them.

This qualitative analysis does not provide a way to quantify something like referential capacity. Indeed, it is not clear what such a measure might consist in. But it does suggest that to utilize the state of a mediating substrate to access the reference afforded by that medium's intrinsic constraints, an interpreting system must do so with respect to the consequences of work. For such a system, functional significance can be assigned to this referential information with respect to work "saved" due to access to information about contextual factors relevant to achieving a preferred target condition. This implies that a system capable of assessing reference must be organized to achieve or maintain a far-from-equilibrium state, such as in a living system.

Although the qualitative form of this analysis may still limit technical applications, such as in molecular biology, cognitive neuroscience, or artificial intelligence, it should nevertheless be sufficient to serve as a framework for the future development of a precise formal theory of reference and functional significance.

References

1. Bateson, G.: Upside-Down Gods and Gregory Bateson's World of Difference. Fordham University Press, New York (1968)
2. Ben-Naim, A.: A Farewell To Entropy: Statistical Thermodynamics Based on Information. World Scientific Publishing Co, Singapore (2008)
3. Boltzmann, L.: The second law of thermodynamics. Populare Schriften, Essay 3, address to a formal meeting of the Imperial Academy of Science, 29 May 1886, reprinted in *Ludwig Boltzmann, Theoretical Physics and Philosophical Problems*, S. G. Brush (Trans.). Boston: Reidel (1874)
4. Brentano, F.: Psychology From an Empirical Standpoint. Routledge & Kegan Paul, London, pp. 88–89 (1874)
5. Brillouin, L.: Science and Information Theory. Academic Press, New York (1962)
6. Chaitin, G.: Algorithmic Information Theory. IBM J. Res. Develop. **21**(350–359), 496 (1977)