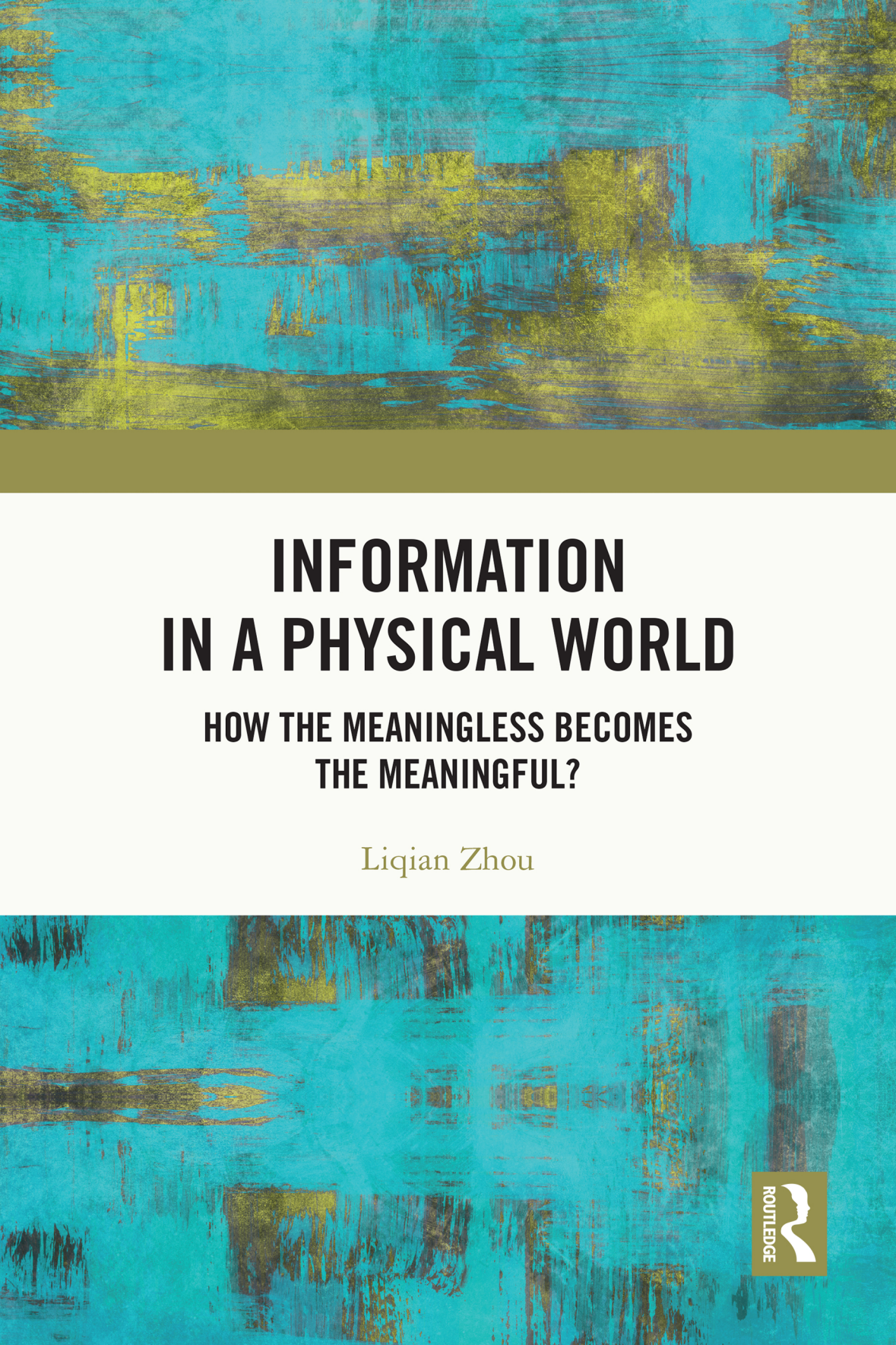




INFORMATION IN A PHYSICAL WORLD

**HOW THE MEANINGLESS BECOMES
THE MEANINGFUL?**

Liqian Zhou



Information in a Physical World

To construct a comprehensive theory of information, meaning and intentionality, the book develops a naturalistic perspective based on Peircean biosemiotics. It re-examines key issues such as referential information, the metaphysics of form, misrepresentation, reference and sense, offering new explanations and insights.

It analyzes the notion of information conceptually, showing that it is not a simple concept, but rather comprises hierarchically nested aspects: Structural, referential and normative. By defining form negatively as a constraint, it offers a naturalistic account of structural information. It then addresses the normativity of referential relations through the concept of interpretation, introducing the notion of operational interpretation to explain the minimal normativity of reference. The book further argues that the three types of reference—symbolic, indexical and iconic—are asymmetrically interdependent: Symbolic reference depends on indexical reference, which in turn depends on iconic reference. This semiotic framework helps to resolve many long-standing problems in the philosophy of language. Finally, the book argues that convention is inherently a semiotic concept. As such, symbolic reference has two aspects of conventionality: The sign vehicle and the referential relation. This semiotic explanation offers a deeper understanding of real-world communication scenarios.

The book will be of interest to researchers and advanced students in philosophy of mind, philosophy of language, philosophy of biology, philosophy of cognitive science, semiotics, biosemiotics and linguistics.

Liqian Zhou is an associate professor of philosophy at the Department of Philosophy, Shanghai Jiao Tong University. His research interests include philosophy of information, philosophy of biology, philosophy of mind and cognition, and biosemiotics.



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Information in a Physical World

How the Meaningless Becomes the
Meaningful?

Liqian Zhou



Routledge

Taylor & Francis Group

LONDON AND NEW YORK

First published 2026
by Routledge
4 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
605 Third Avenue, New York, NY 10158

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2026 Liqian Zhou

The right of Liqian Zhou to be identified as author of this work has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilised in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN: 978-1-041-09725-9 (hbk)

ISBN: 978-1-041-09730-3 (pbk)

ISBN: 978-1-003-65150-5 (ebk)

DOI: [10.4324/9781003651505](https://doi.org/10.4324/9781003651505)

Typeset in Times New Roman
by KnowledgeWorks Global Ltd.

To my mentors:

**Søren Brier, Zhong Cai (蔡仲), Terrence Deacon and
Shuhua Li (李曙华) .**



Taylor & Francis

Taylor & Francis Group

<http://taylorandfrancis.com>

Contents

<i>List of figures and tables</i>	<i>viii</i>
<i>Foreword by Terrence W. Deacon</i>	<i>ix</i>
<i>Acknowledgment</i>	<i>xiii</i>
1 Grounding meaning in information	1
2 A conceptual analysis of information	23
3 Information is physical (negatively)	41
4 Interpretation	59
5 Reference	82
6 Convention	102
<i>References</i>	<i>126</i>
<i>Index</i>	<i>134</i>

Figures and tables

Figures

1.1	A communication system by Shannon (1948)	4
6.1	Signaling systems in a simple signaling game	106
6.2	Two possible signaling systems in the g_1	113
6.3	Two possible signaling systems in the g_2	114
6.4	Two actual possible signaling equilibria in the g_2	115

Tables

6.1	A simple signaling game	105
6.2	Possible combined strategies in a simple signaling game	105
6.3	g_e	110
6.4	g_t	110
6.5	g_b	110
6.6	Stag hunt game, g_1	113
6.7	Possible combined strategies in g_1 as a signaling game	113
6.8	Battle of sexes game, g_2	114
6.9	Possible combined strategies in g_2 as a signaling game	114
6.10	The actual payoff of acts coordination	115
6.11	Actual possible combined strategies in the mutual misunderstanding scenario	115
6.12	The relationship between informational and functional content in actual game	118
6.13	The relationship between informational and functional content for the sender	118
6.14	The relationship between informational and functional content for the receiver	119

Foreword

We no longer find it unusual to see someone using their smartphone to chat face-to-face with a friend in a distant city, even while standing in line waiting to purchase a cup of coffee with that same smartphone. Nor do we think it is magical that your car can guide you to an unfamiliar destination by telling you when to turn left or right down unfamiliar roads. We have even become familiar with the ability of law enforcement to identify a crime suspect by comparing the DNA evidence to data in a public DNA database. None of these examples of modern information technology at work seem remarkable anymore. They are all everyday occurrences in the information age. Curiously, however, if you ask what it is that we call information in these cases we get oddly diverse answers. An engineer might tell you that it's electrical signals. A software designer might tell you that it's bits. A philosopher might tell you that it's meaning. Like the parable of the blind men explaining that an elephant is a hose or tree trunk, we seem unable to grasp the whole of the thing, even when it is ever-present.

It's been three-quarters of a century since Claude Shannon provided his definitive analysis of the mathematics of communication which effectively initiated the dawn of the information age. His insight into the nature of information sparked a revolution that reshaped almost every aspect of life in the 21st century. And yet, ironically, it has also contributed to a troubling ambiguity in our understanding of the nature of information. As many a writer has lamented, the technical precision that Shannon's analysis achieved was in large measure made possible because he was able to bracket any discussion of the referential ("aboutness") and normative ("useful") attributes of information as it is colloquially understood. In a famous *Scientific American* article (that also inspired the foreword to a book version of Shannon's "technical report") Warren Weaver clarifies this special usage as follows:

First off, we have to be clear about the rather strange way in which, in this theory, the word "information" is used; for it has a special sense which, among other things, must not be confused at all with meaning.

And in the introductory paragraph of his "report," Shannon also makes clear that for his purposes this defining property of "information" need not be considered

because “These semantic aspects of communication are irrelevant to the engineering problem.” Weaver further warns that this special use of the word “information” can lead to serious confusion if not carefully distinguished from the common sense understanding of the term.

It is surprising but true that, from the present viewpoint, two messages, one heavily loaded with meaning and the other pure nonsense, can be equivalent as regards to information.

Ironically, the incredible technological impact of this special restricted usage of the term “information” has effectively caused this special meaning to eclipse the colloquial meaning. In the intervening decades the tail has begun to wag the dog, so to speak, with the technical meaning and colloquial meaning often becoming interchanged or used ambiguously as though they are synonymous. This is particularly problematic in more popular treatments of the concept of information, such as in James Gleick’s award-winning book *The Information: A History, A Theory, A Flood*. While recounting in wonderful detail the history and current implications of this revolutionary theory we find no attempt to even consider questions about meaning or use, as though these issues are peripheral.

Of course, the history of philosophical debate over these “peripheral” issues has been raging for millennia and shows no signs of having reached a resolution. In many respects, these debates have been carried out with different terminology over issues of representation, theories of knowledge, the concept of value, the nature of purpose, etc. Since Shannon’s contribution, these philosophical conversations have also begun to incorporate discussions of the relevance of this special interpretation of the information concept. Scholarly efforts to integrate information theory into the philosophy of mind and theories of knowledge have given rise to a number of groundbreaking books on information by such authors as Bar-Hillel and Carnap, Wiener, Dretske, Millikan, Floridi and Skyrms, just to name a few. University departments, international professional societies and professional journals focused on different approaches to information have also grown in numbers over the years. So, 75 years after Shannon, one might reasonably assume that only modest adjustments to our knowledge of information are possible. How could there be any fundamental questions that have not been considered and answered?

The title of this book—*Information in a Physical World*—hints that a fundamental question does indeed remain unanswered. The title cryptically juxtaposes two quite different understandings of the concept of information. On the one hand, bits of information in the form of physical images, texts and numbers are found everywhere in the modern world. This physicality is obvious, but its consequences tend to be relatively minor. On the other hand, the meaning conveyed by these tokens is not any part of their physicality, and yet it is nevertheless causally relevant, often initiating massive changes in societies and their physical surroundings. But this difference in effect seems to violate a common assumption that only physically present properties have causal effects.

The physicality of information is of course taken for granted, at least in one sense. Both brains and computers are physical, and presumably, it is the manipulation of signals, whether chemical or electrical, that constitutes what we call information processing in these systems. One would assume, then, that the physicality of information is well understood. Indeed, most scholars working in the field have agreed: *information is physical*, as Rolf Landauer emphatically proclaimed in an IBM research paper in 1961. It is not some disembodied formal essence. And yet, it turns out that the physicality of information is not so simple. The subtlety of this point becomes more obvious when one reflects on the basis of Landauer's conclusion. The definitive evidence for the physicality of information is provided by the fact that only *erasing* information (for example, from a computer memory) produces an increase in entropy. But this is a strange sort of negative physicality. And at least on the surface, it doesn't seem relevant to an explanation of how meaning can have physical consequences.

Information in a Physical World explores this special sort of physicality that information (in the full semiotic sense of that word) entails. The argument begins by distinguishing three hierarchically related ways of analyzing information: In terms of its structural properties, its referential function and its normative consequences. Shannon's analysis has provided an adequate basis for characterizing the structural aspects of information, while issues of reference and normativity have mostly been debated by philosophers and treated as epiphenomenal in the natural sciences. A major aim of the book is to account for the causal efficacy of these apparently subjective semiotic aspects of information.

This is accomplished by a sort of figure/ground inversion of the accepted view of information in which the referential and normative properties of information provide the primary source of physical efficacy while the structural properties of the information-bearing media are secondary. To explain how the non-intrinsic aspects of information can be a source of physical efficacy the analysis turns to the process of interpretation. This relocates the locus of causal efficacy in the interpretive process rather than in the information medium and shifts the focus of the analysis to the thermodynamic properties of the process of interpretation. Consequently, much of the discussion in later chapters of the book explores the logic of interpretation in its various forms and contexts. At one extreme this involves a discussion of the basic interpretive processes that comprise the normative chemistry of living processes. At the other extreme, this involves an analysis of the complex semiotic infrastructure underlying symbolic thought and language. This shift in focus away from the intrinsic structural properties of information media to the physical properties of interpretive processes demystifies what had once appeared to be a nonphysical causal relation. The apparent non-physicality of semiotic causality turns out to be an artifact of considering information in isolation from the physical interpretive process that invests it with causal relevance.

Although *Information in a Physical World* offers a groundbreaking new perspective on this challenging theoretical problem, it also provides an in-depth

scholarly review of the history of the topic and the many attempts to make sense of the apparent paradox of semiotic causality. In the process of re-analyzing the physicality of information, it also inevitably synthesizes insights drawn from a wide range of disciplines. These include philosophy of mind, information theory, semiotic theory, thermodynamics, origins of life research and linguistics, as well as my own research exploring the nature of information and the origins of life (which he explains in accurate detail and integrates seamlessly into his analysis). Though many of these views concerning the nature of information and the basis of semiotic causality have been influenced by my own research, this new synthesis has opened my eyes to the insight that I will need to ponder for years to come. For others with similar interests, this book provides an indispensable framework for making sense of the new conceptual challenges likely to emerge from advances in molecular biology, cognitive neuroscience and artificial intelligence (AI).

Terrence W. Deacon,
The University of California, Berkeley,
March 2025

Acknowledgment

This book is a reworking of my doctoral dissertation. However, it is entirely newly written, with only a few paragraphs carried over from the original dissertation. The motivation to transform the dissertation into a book came from Terrence Deacon. In 2018, before his visit to China, Terry read the dissertation and encouraged me to publish it. Since the dissertation focuses on my understanding of Terry's theories on life, information and mind, and their application to related philosophical issues, his personal affirmation gave me the confidence to rework it into a book. This work would not have been possible without him. I am deeply grateful to Terry for his lasting inspiration, encouragement and support.

I also owe my thanks to my academic mentors. I am profoundly grateful to my MA supervisor, Shuhua Li, my PhD supervisor, Zhong Cai and my PhD supervisor abroad, Søren Brier. Shuhua provided me with systematic and rigorous training in the philosophy of complexity and systems science, and her pure and honest spirit of scholarship has continued to guide me throughout my career. In late 2011, I decided to continue my philosophical studies at the Department of Philosophy, Nanjing University. Zhong kindly accepted me as his PhD student, despite knowing that our academic interests differed. His patience, encouragement and unwavering support have accompanied me from that point onward. After a year of study at Nanjing University, Zhong encouraged me to pursue further training abroad, which led me to Søren. Søren invited me to Copenhagen Business School, where I spent two years working on my PhD project in the philosophy of information with his supervision. He introduced me to the communities of biosemiotics and information studies and the training he provided continues to benefit me today.

I began my academic career at the Department of Philosophy, Nanjing University, before securing a new position at the Department of Philosophy, Shanghai Jiao Tong University. The dynamic academic environments of these two institutions, enriched by my colleagues, have greatly influenced my work, and I am grateful to all of them. Parts of this book were also based on research projects supported by the China Postdoctoral Science Foundation (no. 18CXZ013) under the project "The Ontological Commitments and Methodological Selection of Contemporary Information Studies" and by the National Social Science Foundation (2017M611789) under the project "Teleosemantics and its Formalization."

Many ideas and draft chapters have been presented at conferences and invited lectures, where I received valuable feedback from colleagues and friends that significantly improved and elaborated the arguments in this book. I am grateful to the following individuals for their insightful comments and support: Qiaoying Lu, Zixia Zhang, Wei Fang, Qianqian Sun, Loki Chen, Jian Wang, Yixin Zhong, Kun Wu, Dongsheng Miao, Xueshan Yan, Chuang Liu, Xiaotao Liu, Qiu Wang, Xingming Hu, Yuanfan Huang and Changsheng Lai, as well as the audiences who heard my lectures and reports: International Society for the Studies of Information, Summit 2017 by Vienna Technology University; Summit 2019 by UC Berkeley; English session in National Biennial Conference of Philosophy of Science by Zhejiang University; 1st Science and Philosophy Workshop for Young Scholars by Nanjing Normal University; Department of Philosophy, Nanjing University; Department of Philosophy, Xi'an Jiao Tong University; Philosophy School, Wuhan University; Department of Philosophy, Chinese University of Science and Technology; School of Philosophy, Fudan University. To the students who attended my seminar, “Frontiers in Philosophy.”

Jeremy Sherman helped me frame the initial outline of the book and gave me detailed comments on [Chapter 2](#), and Mingjun Zhang provided detailed feedback on an early version of [Chapter 6](#). [Chapter 2](#), titled “Structural, Referential, and Normative Information,” has been published in *Information & Culture* 56(3). I am particularly grateful to the anonymous reviewers for their insightful and constructive comments, which helped refresh this chapter. Several ideas and arguments in [Chapter 3](#) were presented to the reading group on self-generative structures, and I would like to thank Weiyi Qiu, Yu Liu, Dong Wang and Mingli Yuan for their contributions. I am grateful to Huijuan Zhou who gave me a comprehensive explanation of the basics of mathematical function and principles of deep learning neuro networks.

We are fortunate to live in the era of generative AI powered by large language models. I use ChatGPT to polish and refine the English in this book, correcting grammar mistakes and typos while enhancing fluency—without altering my writing style. I no longer need to spend a significant amount of money on polishing my English or worry that those helping with the editing might inadvertently alter my intended meaning. Furthermore, the impressive capabilities of generative AI, powered by large language models, offer another compelling example of the universal regularities in the indexical relationships between symbols in language. However, generative AI lacks the ability to interpret symbolic references. The remarkable outputs it generates—sentences, paragraphs, arguments, proofs and so on—are the result of these universal regularities. However, there are no indexical relationships between symbols and objects, meaning that the symbol networks do not correlate with the world. When these models require grounding in the physical world, they encounter the world-modeling problem.

Finally, I owe a debt of gratitude to my family—my parents, my wife and my mother-in-law—who have supported me by taking on responsibilities that I should have shouldered myself.

1 Grounding meaning in information

1.1 Introduction

A gamma-ray burst detected by the *Fermi Gamma-ray Space Telescope* carries information about a stellar system billions of light-years away. The rock strata of a region on Earth, possibly containing fossils, reveal detailed information about the environment of that area hundreds of millions of years ago. The double-helix structure of DNA encodes the genetic information of an organism. A bee's dance conveys the information about the location of nectar. Smoke rising from a beacon signals an invasion to a soldier. The stock index reflects the information of market trends. These examples suggest that information is a natural kind in the universe—ubiquitous and intrinsic to nature. Intuitively, information is *about* something. Furthermore, there are well-established formal theories that characterize information, such as Shannon's mathematical theory of communication and Kolmogorov's algorithmic theory. These frameworks have been widely applied across physics, biology, cognitive science, communication engineering and computer science. Given that information appears to be a natural kind, is inherently *about* something and is rigorously formalized, it is unsurprising that some philosophers have turned to it in search of a naturalistic solution to one of philosophy's most persistent problems: The problem of intentionality.

A distinctive feature of mind and language is their ability to be *about* something. Mental states and linguistic symbols establish intentional relationships with their referents. This *aboutness* is puzzling for several reasons. First, mental states and symbols are obviously not the referents they represent. For example, when I see a Donald Duck doll on my bookshelf, my mind *re-presents* the doll—but the mental state itself is not the doll. Second, mental states can refer to things that no longer exist. When reminiscing, my mind is *about* last night's dinner, even though that event has passed. Third, mental states can be about universals. The statement "Animals are living organisms" refers to the universal concepts of *animals* and *organisms*. Fourth, they can be about fictional entities. The sentence "Monkey King is short" ascribes a property to the character *Monkey King* from the Chinese classic *Journey to the West*, despite his fictional nature. A central task of the philosophy of mind and language is to explain how the mind establishes *aboutness*—how mental states acquire their referents and how language gains its sense and meaning. As

2 Information in a Physical World

Searle (1979, 89) puts it: “The problem of meaning in its most general form is the problem of how ... we can get from the physics to the semantics.”

One of the most influential and promising approaches to this problem over the past four decades has come from naturalism. Most contemporary philosophers identify as naturalists, though the term itself is somewhat ambiguous. Nevertheless, the naturalistic approach to *aboutness* follows a clear agenda. As Neander (2017, 3–4) argues:

...the naturalized semantics project is the attempt to explain the semantic facts and properties of the world, at their most fundamental, in terms of the nonsemantic facts and properties of the world, with an added proviso that is notoriously hard to specify. Suffice it to say here that this proviso is, roughly, that the relevant nonsemantic facts and properties are condoned by the natural sciences.

The naturalistic project argues that *aboutness* or representation has its roots in nature—meaning has a natural origin.

Paul Grice (1957) distinguishes between *natural meaning* and *non-natural meaning*. Natural meaning, as the term suggests, exists in nature. Typical examples include smoke indicating fire, dark clouds signaling impending rain and the number of rings in a tree stump revealing the tree’s age. Some philosophers believe that natural meaning serves as the foundation for non-natural meaning. Since information appears to be a natural kind—well-defined and well-structured—many argue that it is a strong candidate for natural meaning. They propose that information underlies meaning and intentionality, offering a naturalistic explanation. Since information exists inherently in nature, the challenge of intentionality then becomes explaining how living organisms utilize informational relationships to represent things relevant to their needs.

This approach seems promising and has led to productive research over the past few decades. However, when we examine how these philosophers define *information*, we find significant variation in their explanations. This is not a new issue in information studies. While the concept of information plays a unifying role across disciplines such as physics, biology, cognitive science and social science, its interpretation varies considerably across fields. It is true that we have good formal theories of information and which have advanced the communication, computation, information technologies and facilities to be a necessary commodity in our daily life. These successes come at a cost—the loss of information’s very meaning.

The engineering conception of information focuses solely on quantitative analysis, disregarding the content and pragmatic use of information—features that are essential to our understanding of it. Intuitively, information is something carried by signals *about* something *for* some use. However, formal theories primarily address the *quantity* of signals, a narrow aspect of information’s physical nature. This limitation makes them insufficient for explaining how physical signals convey abstract content, how signals come to carry specific meanings and how abstract content can have physical consequences. As a result, current information theories are inadequate for grounding a theory of meaning or intentionality.

Nevertheless, this does not mean that the project of grounding meaning in information is doomed to fail. On the contrary, I share the view that this approach remains promising—provided we develop a more comprehensive theory of information. Such a theory must account for the *structural*, *referential* and *normative* aspects of information, thereby laying the foundation for a convincing naturalistic account of reference and meaning. This is the task of this book.

To construct a comprehensive theory of information, this book will undertake the following: (1) Conceptual Analysis. Demonstrating that information consists of three nested aspects: Structural, referential and normative. Existing theories have largely focused only on the structural aspect. (2) Critical review of naturalistic theories of meaning. Arguing that the difficulties faced by current naturalistic approaches stem from an inadequate understanding of information. (3) Theory Construction. Drawing on Peircean semiotics, the book proposes that information should be understood as a triadic phenomenon: Carried by signals *about* something *for* some use. Information is inherently relational and *negative*. This negative perspective clarifies the relationships between information, thermodynamics, interpretation and evolution. Through interpretation, informational relationships form nested referential structures: the *iconic* depends on the *indexical*, which in turn depends on the *symbolic*. (4) Application. (a) Reformulating the *symbol grounding problem as the symbol ungrounding and regrounding problems within this framework*, shifting the focus to how symbolic relationships can be ungrounded without losing their referential power. (b) Providing a plausible foundation for a naturalistic approach to intentionality, reference and meaning. The theory offers solutions to the problems of misinformation and content determination.

The purpose of this introductory chapter is to clearly define the specific problems this book seeks to address. By critically reviewing previous efforts, it establishes the need for a new approach and outlines the book's structure as a guide for the arguments that follow.

1.2 Shannon's mathematical theory of communication

The idea that Shannon's mathematical theory of communication is incapable of accounting for the meaningful aspect of information is not new. In his seminal paper, Shannon explicitly states, "Frequently the messages have meaning; that is, they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem" (Shannon 1948/1964, 31). In an interview, Robert Fano further comments on the theory, noting that,

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)

4 Information in a Physical World

Furthermore, it is precisely this ignorance of the “aboutness” aspect of information that has contributed to the great success of Shannon’s theory. As John Collier (2003) comments, “The great tragedy of formal information theory is that its very expressive power is gained through abstraction away from the very thing that it has been designed to describe.” As a result, ever since the early days of information and computation science, researchers have been on a mission to move beyond Shannon and develop a theory that can characterize the very essence of what defines information—namely, reference and meaning. Two types of work have been undertaken: One focusing on articulating the relationship between Shannon’s theory and the “aboutness” aspect of information, and the other searching for a way to characterize this “aboutness.”

To better understand these endeavors, it is necessary to first understand the theory from which they departure, namely, Shannon’s formal theory. Almost all later theories of informational content begin with Shannon’s theory, either criticizing its deficiencies or following its logic and ideas. In this section, I briefly introduce the theory based on Shannon (1948) and Lombardi et al. (2016).

Shannon defines information within the context of a communication system. According to Shannon (1948), a communication system consists of five components: The information source, transmitter, channel, receiver and destination. These are defined as follows: (1) An information source produces a message (or sequence of messages) to be communicated to the destination. (2) A transmitter transforms the message(s) into signals suitable for transmission over the channel. (3) The channel is the medium through which the signals are transmitted from the transmitter to the receiver. (4) The receiver performs the inverse operation of the transmitter, reconstructing the message(s) from the received signals. (5) The destination is the entity or location to which the message(s) is intended to be sent. The communication system is typically formulated as follows:

Since the message is the meaningful part, and is irrelevant to the engineering problem, Shannon focuses solely on the processes between the transmitter and receiver—specifically, the transmission of signals. His goal is to provide a formal way to measure the quantity of information carried by signals and the capacity of the channel. The amount of information carried by a signal is not determined by

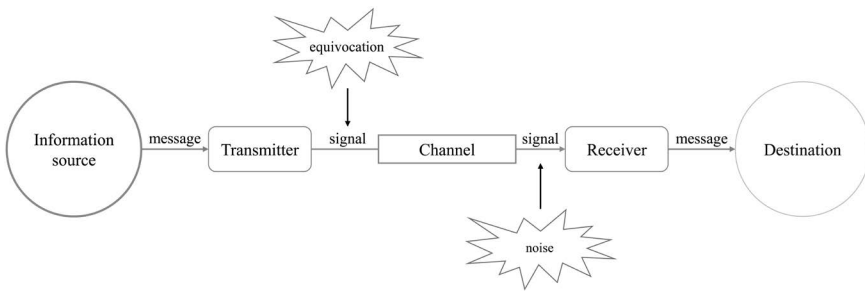


Figure 1.1 A communication system by Shannon (1948)

the intrinsic properties of the signal itself, but by the set of possible signals it was chosen from. In his terms, Shannon defines information as the reduction of uncertainty, which can be formalized using probability theory.

Suppose an actual signal is selected from a set with two equally likely signals. In this case, the quantity of information carried by the signal can be measured by the amount of uncertainty it resolves. Using the logarithmic function with base 2, the signal carries 1 bit of information. Now, suppose the information source S is a system with a range of states $s_1, s_2, \dots, s_n$ and the probability of each state occurring is equal. The quantity of information $I(s_i)$ generated at the source by the occurrence of state s_i can be defined as:

$$I(s_i) = \log \left(\frac{1}{p(s_i)} \right) = -\log p(s_i)$$

Since S produces a range of states, then the informational entropy of S is,

$$H(S) = \sum_{i=1}^n p(s_i) \log \left(\frac{1}{p(s_i)} \right) = -\sum_{i=1}^n p(s_i) \log p(s_i)$$

The destination is also a system with a range of possible states $d_1, d_2, \dots, d_m$. The information generated by the occurrence of a state d_j , $I(d_j)$ and the informational entropy of D can also be measured in the same way.

The amount of information transmitted by a signal from the information source to the destination may not necessarily equal either of these two quantities. The amount of information carried by a signal is measured as mutual information. Mutual information quantifies the average amount of information generated at S and received at D . The amount of mutual information depends not only on the information generated at S but also on noise and equivocation. In an ideal, noiseless channel, the amount of information received at D , $H(D)$, would equal the amount of information generated at S , $H(S)$. However, in actual communication systems, the channel is not noiseless. Noise and equivocation can distort the amount of information transmitted by the signals.

Technically, noise N represents the average amount of information received at D but not generated at S , while equivocation E represents the average amount of information generated at S but not received at D (Lombardi et al. 2016). Thus, mutual information can be defined as:

$$H(S; D) = H(S) - E = H(D) - N$$

Mutual information measures the degree of dependence between the destination D and the source S . If S and D are independent, then the equivocation E and noise N are maximized. In this case, $H(S) = E$ and $H(D) = N$, meaning there is no mutual information between S and D , that is, $H(S; D) = 0$. If there is dependence between S and D , then $H(S) > E$ and $H(D) > N$. The greater the dependence between S

and D , the smaller the values of E and N . The maximum mutual information also serves to measure the channel capacity, which reflects the reliability and maximum throughput of the channel—the maximum amount of information the channel can reliably transmit.

The next issue is how to calculate the quantity of mutual information. Since information is defined as the reduction in uncertainty, mutual information can be understood as the entropy resulting from changes in the probabilities of events occurring at D , given that events have occurred at S . Given that events occur at S , the probability of events occurring at D is $p(d_j|s_i)$. This is determined by the joint probability of both s_i and d_j occurring, $p(s_i, d_j)$, divided by the probability of the occurrence of d_j , $p(d_j)$. Essentially, this measures how much $p(s_i, d_j)$ is proportional to $p(d_j)$.

$$p(d_j | s_i) = \frac{p(d_j, s_i)}{p(d_j)}$$

Accordingly, we can calculate $I(d_j|s_i)$, the information generated by d_j given that s_i has occurred and $H(S)$. The difference between the entropy at the destination, $H(D)$, and the conditional entropy at the destination given the source, $H(D|S)$, represents the mutual information transmitted from S to D through the signal.

$$MI(S; D) = H(D) - H(D|S)$$

Expand the formula, we get,

$$MI(D, S) = \sum_{i=1}^n \sum_{j=1}^m p(d_j, s_i) \log \left(\frac{p(d_j, s_i)}{p(d_j)p(s_i)} \right)$$

Given that an event d_j occurs at D , the amount of information received at S is the same. This reciprocal relationship is why it is referred to as mutual information.

Now, with the fundamentals of Shannon's information theory in place, we can critically review the efforts to extend beyond this theory and recover the aboutness or meaningful aspect of information—an aspect that Shannon intentionally disregarded in his original framework.

1.3 Theories of informational content

1.3.1 Bar-Hillel and Carnap

Although Shannon and his colleagues clearly warned that his formal theory has nothing to do with the meaningful aspect of information, the theory has been almost immediately embraced as a general theory of information covering both the formal and meaningful aspect of information since its inception. Therefore, in order to provide a comprehensive theory of information, we must first clarify

this confusion. This task was taken up in one of the earliest efforts by Yehoshua Bar-Hillel and Rudolf Carnap (Carnap and Bar-Hillel 1952a; 1952b; Bar-Hillel 1953/1955; Bar-Hillel and Carnap 1953a, 1953b; Bar-Hillel 1955).

Bar-Hillel and Carnap first carefully address the confusion surrounding the relationship between Shannon's information theory and the aboutness/meaningful aspect of information, or informational content. While Shannon and others were clear that his theory is a quantitative measure of information that does not concern content, "it is psychologically almost impossible not to make the shift from the one sense of information, for which this argument is indeed plausible, that is, information = signal sequence, to the other sense, information = what is expressed by the signal sequence, for which the argument loses all its persuasiveness" (Bar-Hillel 1955, 94). As introduced in the previous section, Shannon's method for measuring the quantity of information involves the possibility of choosing a message from a set. The rarer a message is chosen, the more information it carries. Since the messages are represented by signals or symbol sequences, Shannon essentially measures the rarity of these signals.

But it must be perfectly clear that *there is no logical connection whatsoever between these two measures, i.e. the amount of (semantic) information conveyed by a statement and the measure of rarity of kinds of symbol sequences*, even if these symbol sequences are typographically identical with this statement. The event of transmission of a certain statement and the event expressed by this statement are, in general, entirely different events, and the logical probabilities assigned to these events, relative to certain evidence, will be as different as will be the frequencies with which events of these kinds will occur relative to certain reference class."

(emphasized by the original paper, Bar-Hillel 1955, 95–96)

Nevertheless, they argue that although Shannon's formulas pertain to signal sequences and have nothing to do with the meaningful aspect of information, they still reflect general formulas about information (Carnap and Bar-Hillel 1952; Bar-Hillel and Carnap 1953). Rather than measuring the quantity of information by the frequency of signals, Bar-Hillel and Carnap use the inductive probability of statements within the set of all possible descriptions of the universe's states. The amount of informational content carried by a statement is measured by the number of statements it excludes. The more statements a given statement excludes, the more content it carries. Formally, the method to measure the quantity of informational content is identical to the formulas in Shannon's theory, and I will not delve into the details here.

This project was an important attempt but has several unresolved issues. First, there are logical results that are counterintuitive, as Floridi (2004) has pointed out and as Bar-Hillel and Carnap (1952) recognized from the outset. A tautology carries zero content because it is self-contained and thus excludes no other statement. In contrast, a self-contradictory sentence carries the most inclusive information because it excludes every other statement due to its inconsistency. These results

are counterintuitive. Many proofs in mathematics can be seen as tautologies because their inferences are deductive. However, empirically, when we discover a new proof, we do gain information. Regarding the second result, we generally do not consider self-contradictory sentences to carry any information. Floridi calls the second logical difficulty the Bar-Hillel-Carnap Paradox.

Second, some of the assumptions in the theory are questionable. The formal theory employed by the project is a first-order propositional logic-based induction theory by Carnap. The basic elements of the theory are propositions. Bar-Hillel and Carnap's theory assumes that the number of propositions describing all possible states of the universe is finite. This assumption is dubious because it is possible to offer almost infinite semantic descriptions of the universe's states, as the combinations of properties can be infinite. As a result, any signal could theoretically carry an infinite amount of information without a proper reference class, creating an infinite paradox in the theory. Nevertheless, Bar-Hillel and Carnap's endeavor illustrates that the relationship between Shannon's theory and informational content is more complex than early researchers initially thought.

1.3.2 *Dretske's semantic theory of information*

Dretske's semantic theory of information is another influential approach, particularly in the effort to naturalize meaning and intentionality (Dretske 1981; 1988). Dretske defines informational content as follows:

Informational content: A signal r carries the information that s is F = The conditional probability of s 's being F , given r (and k), is 1 (but, given k alone, less than 1).

(Dretske 1981, 65)

In this definition, if an event (signal) carries informational content that s is F , given the occurrence of the event, s 's being F must occur. In other words, for Dretske, there is a necessary relationship between the two variables if one carries informational content about the other. This necessary relationship should be a nomic (lawful) regularity between the event types that the two variables represent: "a regularity which *nomically precludes* r 's occurrence when s is not F " (Dretske 1988, 245). Thus, s 's being F lawfully depends on r . This implies that a mere correlation between two types of variables, even if perfect, is insufficient to establish an informational relationship (Dretske 1988, 73).

The " k " in the definition refers to the background knowledge of the receiver, meaning what the receiver already knows about the possible states at the source. Dretske argues that "*how much* information a signal carries, and hence *what* information it carries, depends on what the potential receiver already knows about the various possibilities that exist at the source" (Dretske 1988, 79). This introduces the relative aspect of information: Informational content is relative to the background knowledge of potential receivers. However, this claim seems to conflict

with Dretske's assertion that information is an objective commodity. He argues that information, like weight and height, is objective but relative.

An important implication of this definition is the principle of nested relations of informational content: "The information that t is G is nested in s 's being $F = s$'s being F carries the information that t is G " (Dretske 1988, 71). A piece of information can be nested in a state of affairs in two senses: Analytically and nomically. For example, "John is male" is analytically nested in "John is a bachelor" because the definition of the latter implies the former ("A bachelor is an unmarried man"). Nominally nested relations are guaranteed by natural laws. Since a signal is always an event being employed as such, it carries a variety of different pieces of information nomically. "Signals, it seems, are pregnant with information" (Dretske 1988, 73), and "no single piece of information is entitled to *the* status of the informational content of information" (Dretske 1988, 72). As a result, we should distinguish between the informational content a signal carries and the informational content the receiver is interested in and successfully extracts.

Dretske's theory is influential in the project of naturalizing meaning with information; however, it faces several criticisms. The first critique is the strength problem (Millikan 2004, 32; Stegmann 2015). The theory requires that the conditional probability should be 1, which is too stringent. On the one hand, statistical frequency, rather than natural necessity, is what living organisms typically depend on when employing natural information for specific purposes. For example, when a rabbit sees a warning signal about the presence of a predator, the probability that a predator will appear, given the signal, is not 1; it's much less than 1. Nevertheless, this is sufficient for the rabbit—it's a matter of life and death. Even if the signal only truthfully carries information 1 out of 100 times, it is still valuable. On the other hand, an informational relationship does not require a perfect nomic regularity. A mere correlation, even if imperfect, can suffice. For instance, a rabbit may interpret a bush swaying as a signal indicating a predator is behind it. If statistical frequency were the only basis for an informational relationship, then a question arises: How strong must the frequency be? Ruth Millikan, whose work we will discuss in the next section, addresses this issue. Since nomic regularities (natural laws) are not about particulars, there is no law about a bushing being swinging and the appearance of a predator. The correlation between these two events is statistical. Even natural laws, which can account for statistical regularities, do not guarantee that the conditional probability will be 1—this is evident in fields like thermodynamics and quantum physics.

By combining the background knowledge of potential receivers with the principle of nested information, we can distinguish between the information that a receiver is interested in and successfully extracts from signals and the information actually nested in the signals. This combination presents a tension. On one hand, the nested relationship of informational content is consistent with Dretske's claim that information is objective. On the other hand, the amount and type of informational content a signal carries depend, to some extent, on the background knowledge of the potential receiver about the possibilities at the source. It is true that information is relative does not make it subjective. However, the background knowledge is

included in the definition of informational content, as we see above. This implies a divergence between the information nested in a state of affairs and the information relative to a particular receiver. As a result, Dretske's theory implicitly suggests two kinds of information, even though it does not clearly distinguish them.

A third critique of Dretske's theory is that the principle of nested relations of informational content is counterintuitive and may lead to the infinite paradox. The principle seems counterintuitive because in deductive inference, we still gain information when deriving a conclusion, even though it is implied by the premises. For example, Fermat's Last Theorem is implied in the natural number system, yet Andrew Wiles's proof of it in 1994 still provided valuable information. Dretske might argue against this by emphasizing the distinction between the objective information embedded in an event and the information successfully extracted by the receiver. A receiver who receives a signal does not automatically extract all the information embedded in it. Information does not equal knowledge. The information gained by a proof is the information extracted by the receiver, considering their background knowledge. However, this leads to a potential paradox. If any sentence or event contains an infinite amount of information because it implies everything that contradicts it, then even a simple statement like "Lucky is a dog" implies an infinite number of things ("Lucky is not a cat," "Lucky is not a cow," "Lucky is not a tree," and so on). This suggests that every event or statement could contain infinite information, which is problematic.

1.3.3 Floridi's theory of strongly semantic information

Due to the logical dilemma inherent in the Bar-Hillel and Carnap theory, [Floridi \(2004\)](#) proposes an alternative theory of semantic information that avoids the Bar-Hillel-Carnap Paradox. He argues that the root of the paradox lies in the probability-based approach, which treats alethic values as irrelevant to the quantitative analysis of semantic information. To resolve this, Floridi develops a theory of strongly semantic information based on a calculus of truth-values and semantic discrepancy in relation to a given situation. In his theory, alethic value is intrinsic to the definition of semantic information. Informational content is defined in an ideal context, extensional and a priori, as a function of the positive or negative degree of "semantic distance" or deviation from a fixed point or origin—represented by the given situation to which it is supposed to refer ([Floridi 2004](#), 205). The quantity of informational content in a message is measured by the degree of discrepancy: The smaller the discrepancy, the greater the informational content of the message.¹

Now, let's examine how Floridi's theory addresses the Bar-Hillel-Carnap Paradox. The short answer is that Floridi's theory avoids the paradox because it measures the quantity of informational content differently from the Bar-Hillel and Carnap approach.

On the one hand, $\text{CONT}(\sigma)$ refers to *the quantity of semantic information* that can be attributed to σ a priori, on the basis of its probability distributions and independently of the state in which the system under analysis actually is

(context of total ignorance). On the other hand, $\iota^*(\sigma)$ refers to the *quantity of semantic information* that can be attributed to σ still a priori but in a context which is presupposed to be of “localised omniscience” in the game-theoretical sense of perfect and complete information about the system, on the basis of σ ’ alethic value and its degree of discrepancy, relative to a fixed state w of the system under analysis.

(Floridi 2004, 212–214)

The paradox is a logical consequence of Bar-Hillel and Carnap’s theory, and it is expected that a different theoretical system would not lead to the same paradox.

We can also calculate the quantity of informational content for tautologies and self-contradictory sentences according to Floridi’s theory and observe that the results are not counterintuitive. Within this framework, a tautology is consistent with any situation, making it contingently true but vacuous. In the case of self-contradictory sentences, their content is self-referential. The degree of discrepancy for these sentences is measured by how much their content diverges from itself. Because self-contradictory sentences imply maximal discrepancy, they contain no informational content. Another way to understand this is that, being entirely false, self-contradictory sentences are inconsistent with any situation, thus carrying no informational content. Therefore, Floridi’s theory aligns with common-sense intuition and does not result in the Bar-Hillel-Carnap Paradox. Floridi further explains that the reason Bar-Hillel and Carnap’s theory faces the paradox is that their theory measures the quantity of data, not semantic information. This measure depends not on the content of the data but on the number of propositions that the data excludes (Floridi 2004, 214–215).

However, Floridi’s theory of informational content faces several challenges. First, there is an internal tension in the theory (Fresco and Michael 2016). On one hand, information is considered truthful data, but on the other hand, the theory defines information through semantic divergence. This allows for measuring both positive and negative informativeness of a message. Negative informativeness refers to the degree of error. While it is true that we can extract information from even a false message, Floridi asserts that information is inherently truthful, encapsulating truth in itself. This leads to a self-contradiction—how can a false message carry truth? This issue has been widely criticized, with many authors challenging what has been called the “Veridicality Thesis” (Long 2014; Ferguson 2015; Fresco and Michael 2016; Fresco et al. 2017).

A more serious challenge to Floridi’s theory is that it abandons Shannon’s original insight that the quantity of information carried by a signal is determined not by the signal itself but by other possible signals. According to Floridi’s theory, the quantity of informational content of a message is independent of other possible messages. Adraans (2010) argues that Shannon’s formulation represents one of the fundamental achievements of modern science and is consistent with the evidential confirmation in scientific practice. In contrast, Adraans suggests, “Floridi’s philosophy of information is more a reprise of classical epistemology that only pays lip service to information theory but fails to address the important central questions

of the philosophy of information” (Adraans 2010, 41). This oversight of Shannon’s insights results in another deficiency in Floridi’s theory: It cannot address information processing in cognition, since “the same information-processing cognitive mechanisms process information indiscriminately regardless of whether that information is true or false” (Fresco and Michael 2016).

1.3.4 *Vector theories of informational content*

A more promising theory of informational content has been developing in recent years, proposed by several authors (Skyrms 2010a, 2010b; Birch 2014; Scarantino 2015; Shea et al. 2018; Skyrms and Barrett 2019; Isaac 2019). This theory begins with Shannon’s formula but introduces a crucial difference: “informational content must be a vector” (Skyrms 2010a, 2010b, 161). In Shannon’s theory, when measuring the quantity of information a signal carries about a state, only the change in the probability of that specific state is considered. In contrast, the vector approach indicates that informational content should account for how a signal changes the probabilities of each state at the source. Informational content, therefore, has multiple components, one for each possible state. Each component measures the change in the probability of that particular state. For example, consider a source that can be in one of n states ($s_1, s_2, \dots, s_n$). Given a signal m , the informational content of the signal is:

$$\left\langle \log \left[\frac{p(s_1 | m)}{p(s_1)} \right], \log \left[\frac{p(s_2 | m)}{p(s_2)} \right], \dots, \log \left[\frac{p(s_n | m)}{p(s_n)} \right] \right\rangle$$

Scarantino (2015) further develops Skyrms’ theory by distinguishing between two types of informational content: Incremental natural information and the degree of overall support. Skyrms’ theory addresses only the former. Scarantino argues that incremental natural information alone is insufficient to fully characterize informational content. First, simply listing the quantity of information carried by a signal does not provide a complete understanding of its informational content. “Informative signals do not tell us just *how much* probabilities have changed; they also tell us what are *the states of affairs* that had their probabilities changed” (Scarantino 2015, 429). According to this theory, two signals that change the same amount of probability may refer to entirely different states of affairs. The second flaw in Skyrms’ theory is that it omits posterior probabilities from its definition of informational content. In addition to changes in probabilities, we also need to know the overall probabilities of the states of affairs after receiving the signal.

Based on this analysis, Scarantino (2015, 429) proposes that a complete description of a signal’s informational content should include two components: The changes in probabilities, which underlie the transmission of incremental information, and the final overall support conferred by the signal, given the background data. For a source that can be in one of n states ($s_1, s_2, \dots, s_n$), with a signal m and background data d already acquired by potential receivers, the informational content of the signal can be defined as follows:

Incremental information:

$$\left\langle \log \left[\frac{p(s_1 | m \& d)}{p(s_1 | d)} \right], \log \left[\frac{p(s_2 | m \& d)}{p(s_2 | d)} \right], \dots, \log \left[\frac{p(s_n | m \& d)}{p(s_n | d)} \right] \right\rangle$$

Degree of overall support:

$$\langle p(s_1 | m \& d), p(s_2 | m \& d), \dots, p(s_n | m \& d) \rangle$$

Scarantino calls his theory of informational content the Probabilistic Difference Maker Theory. According to this theory, given a signal with respect to a source, we can learn from it: (1) which states of affairs have had their probabilities changed, (2) how much the posterior probabilities have changed and (3) what the posterior probabilities are (Scarantino 2015, 429). I believe this theory is comprehensive because it aligns with both our intuitions about information and avoids the issues faced by other theories.

An intuition captured by Shannon's theory is that information reduces a receiver's uncertainty.² However, Shannon's theory does not account for the fact that information is about something. A more complete formulation of this intuition is that information reduces a receiver's uncertainty about something. Simply put, information tells a receiver something they didn't know before. In the Probabilistic Difference Maker Theory, the incremental information captures the reduction in uncertainty, while the degree of overall support reflects the receiver's posterior epistemic state regarding the source. This aligns with our common-sense understanding of information.

Although the Probabilistic Difference Maker Theory uses Shannon's general formulas, much like Bar-Hillel and Carnap's theory, it does not result in the Bar-Hillel-Carnap Paradox. As Fresco and Michael (2016) have argued, this paradox is superficial if we distinguish between informational content and informativeness. Informativeness here refers to incremental information. It is logically consistent that a self-contradictory sentence implies the maximum quantity of information because a contradiction excludes all possible alternatives. However, information, in this case, refers to what a receiver successfully extracts from a signal—the degree of overall support. Since any self-contradictory sentence is necessarily false, its degree of overall support is 0. Another paradox encountered by other theories of semantic information is the infinite paradox, although it arises for different reasons. The Probabilistic Difference Maker Theory avoids this paradox because it is relative to background data and does not base informational relationships on nomic regularities; correlation alone suffices (Scarantino 2015).

I have now reviewed the main theories of informational content. These theories generally assume that information is meaningful and attempt to characterize the meaningful aspect of information. This project parallels another important endeavor in information studies: Explaining how meaningful data is possible in the first place. The question of how to characterize the measurement of informational

content is, of course, relevant to the project of grounding meaning in information, as many authors have argued (e.g., [Dretske 1981](#); [Stegmann 2015](#); [Scarantino 2015](#); [Isaac 2019](#)). However, the second project is more directly relevant here. We can frame the issues this project addresses through Floridi's formulation of the open problems in the philosophy of information ([Floridi 2011](#)). The most pertinent problems are the semantic problems: How can data acquire meaning? How can meaningful data acquire truth value? Can information explain meaning? To ground meaning or intentionality in information, researchers must solve two main problems: (1) How does information acquire its meaningful or aboutness aspect? (2) Can informational relationships ground meaning? In the next section, we will see that many philosophers have made significant efforts to address these questions.

1.4 From natural information to conventional meaning

Philosophers who attempt to ground meaning in information often begin with a distinction made by Paul [Grice \(1957\)](#): Natural meaning versus non-natural meaning (representation/intentionality). A key difference between these types of meaning is that natural meaning cannot be false, whereas non-natural meaning can be. In other words, when the intended purpose of a sign is not achieved under normal conditions, the non-natural meaning is considered false, whereas the natural meaning remains true. This distinction implies that non-natural meaning has a normative aspect, while natural meaning does not. For instance, if smoke appears without fire, it is not a natural sign of fire.

Some philosophers argue that the distinction between natural and non-natural meaning offers a way to naturalize meaning, grounding non-natural meaning in the natural. To achieve this, two main questions must be answered: First, what exactly is natural meaning, and what is its place in nature? Second, how does the normative aspect of non-natural meaning arise from the natural? In this section, I will briefly review the main approaches to these questions.

1.4.1 Stampe's causal theory of representation

[Stampe \(1977\)](#) offers an early proposal, suggesting that causal correlations in nature provide natural meaning. However, causal relations themselves do not possess a normative aspect; they cannot be accurate or inaccurate, correct or incorrect. It is through the interpretation and intention of a speaker that a state of affairs is identified as the one that would cause the production of a representation, assuming certain causal conditions govern that process. According to Stampe: "[T]o know that *R* represents *O* as being *F* is to know that if *R* is produced under fidelity conditions, then it would be the case that the fact that *O* is *F* would explain (etc.) the occurrence of *R*" ([Stampe 1977](#), 50). This means that, given fidelity conditions under which one state of affairs causes the production of another (as a representation), the representation is expected to convey information about the original state

of affairs. Stampe uses the concept of “function” to explain the normative aspect of representation. A function is supposed to achieve a specific end, but that does not mean it always succeeds in doing so. A function may fail. Similarly, a representation can be false.

Later philosophers have critiqued Stampe’s account for several reasons. First, the theory suggests that a speaker must identify the fidelity conditions of a representation when using it. This does not align with real-world circumstances, where speakers typically do not know these conditions. For instance, when a rabbit perceives the rustling of a bush as a sign of a possible predator, it does not consciously identify the fidelity conditions of that representation (and, in fact, it lacks the competence to do so). Second, Stampe’s reliance on the speaker’s interpretation and intention to explain the normative aspect of representation raises further concerns. While it is true that individuals can assign purposes to objects, as in the case of artifacts designed to perform specific functions, this approach faces challenges. Empirically, most representations—whether linguistic or animal signs—do not have identifiable designers. Moreover, individuals’ intentions and interpretations of representations often diverge from the meanings intended by others. If interpretation is the source of the normativity of representation, additional explanations are needed to account for these discrepancies. Furthermore, the theory must clarify how a hearer can accurately interpret the speaker’s intentions when receiving a representation. Third, causal relations are too rigid to account for many forms of representation. As discussed in the critique of Dretske’s theory, statistical correlations seem sufficient for establishing representational relations. Additionally, many conventional representations are not determined by causal relations at all. A representation may not have any physical connection to the object it represents. In fiction, the represented objects may not even exist physically.

1.4.2 Dretske’s informational semantics

Although later philosophers abandoned causal relations as the candidate for natural meaning, they retained the idea of function as an explanation for the normative aspect of non-natural meaning and sought alternative ways to explain the concept of function. A new candidate for natural meaning that philosophers turned to is information. To explain the normativity of representation, they appeal to biological functions. These two threads together form a family of theories known as teleosemantics. In this context, to provide a naturalistic theory of intentionality, meaning, or representation, philosophers must tackle two key tasks: (1) Explaining what natural information is, and (2) explaining how the normativity of representation emerges from natural information. Specifically, they must address questions such as: What kind of relations can count as natural information? Does an informational relationship need to be held for some reason? Not all natural information is used as non-natural meaning or representation; thus, how is it determined? This issue is called the content determination problem. Furthermore, how can normativity arise

apart from individual intention and interpretation? This is known as the misrepresentation problem.

Building on Grice's distinction and Stampe's causal theory of representation, Dretske (1981; 1986; 1988) refines the classification of representational systems into three types. The first type, symbols, lacks intrinsic representational power. Instead, its power is derived from its creators or users. Symbols are assigned indicator functions that they cannot perform intrinsically. The second type of representational system acquires its representational power from its constituent elements, which are considered natural information. As defined in the previous section, natural information plays the role of natural meaning. For Dretske, natural information exists in nature and does not inherently function as a representation. It is only when users employ this natural information that they assign it a representational function. The difference between these two systems is that in the first type, the function comes first, while in the second type, the power of the elements to indicate comes first (Dretske 1988, 61). For both types, their normativity derives from their creators or users, depending on the interests, purposes and capacities of other interpretive systems. Dretske refers to these as conventional systems of representation. The most interesting type of representational systems, according to Dretske, are natural systems of representation. "Natural systems of representation... are those which have *their own* intrinsic indicator functions, functions that derive from the way the indicators are developed and used *by the system of which they are a part*" (Dretske 1988, 62). Their function of indication is independent of other systems. So, where do natural systems acquire their function of indication? Rather than appealing to evolution by natural selection, Dretske argues that it comes from the development of the individual (Dretske 1988, 95–106).

As mentioned in the previous section, Dretske's formulation of natural information based on nomic regularities is too strong to be universally applicable, and no natural laws specifically address individuals. Furthermore, Dretske's natural information is context-free, neglecting the channel conditions that are also crucial for a sign to carry information effectively. This creates a problem: How should we define the context in which natural information is situated? Millikan (2004, 37–38, 2017, 143–144) calls this the reference class problem. While Dretske appeals to the development of the individual as the source of the normativity of representation, as Stampe did, further explanation is needed to clarify how normativity emerges from this development.

1.4.3 *Information and teleosemantics*

Ruth Millikan (1984; 1989; 2004; 2017) offers a different approach to the issues surrounding meaning and representation. Like Dretske, she believes that non-natural meaning carried by intentional signs depends on natural information, which is carried by locally recurrent natural signs (2004) or infosigns (2017), but she addresses the shortcomings of Dretske's account. Millikan proposes a theory of *infosigns* that resolves many of the problems Dretske's theory faces. Her approach is more organism-friendly. To understand what natural information is, she suggests

that we consider it from the perspective of organisms that use it to perform specific functions (Millikan 2004, 44). For an organism to use natural signs, it must be able to detect the information-bearing signal, identify the presence of an information channel and understand the mapping function that translates the natural signs to their signifieds (Millikan 2004, 44; 2017, 147). Thus, a natural information sign, or infosign, is one that can serve as a sign for organisms (Millikan 2017, 144). For an event *A* to be an infosign of another event *B*, *A* must recurrently correlate with *B* for a reason in a specific context, grounded in a local domain.

Different from Stampe's and Dretske's accounts, Millikan's theory argues that a statistical correlation is sufficient for establishing a natural informational relationship. As long as the correlation between *A* and *B* persists for a reason, meaning there is a non-accidental dependency of *B* on *A* in nature, *A* is an infosign of *B*. Since the theory only requires the correlations to be statistical, they can be about individuals as long as they are well-grounded in local domains (Millikan 2004, 42). Moreover, the correlations are between types of states of affairs. "Correlations between states of affairs that underlie infosigns must be such that one is determined as in a certain historical relation to the other" (Millikan 2017, 149). Therefore, for an event to be an infosign of another event, the correlation between them should recur, and the two particular events must be members of two corresponding types of events. The strength of the correlation can vary, being either strong or weak. Millikan argues that this statistical consideration of natural information is more realistic.

Unlike Dretske's theory, which only considers context-free information and overlooks channel conditions, Millikan's theory insists that correlations, as natural information, always exist in local domains. They are defined relative to a reference class. In other words, to understand an infosign, we must also consider the channel conditions that support it. These channel conditions constitute the reference classes of infosigns, and the reference class of an infosign should not be arbitrary. For the notion of an infosign to explain why an organism might use it as an indicator of its signified with some success, it must have its footing in nature (Millikan 2004, 38–39). That is, the reference class of an infosign is determined by the local context where the infosign and its signified recurrently correlate. By considering the reference class, Millikan's theory avoids the challenges faced by Dretske's context-free theory and is more user-friendly.

Furthermore, Millikan provides a detailed explanation of the corresponding relationship between infosigns and their signified. In a natural informational relationship, both the infosign and its signified are structured world affairs. "Their structure determines their meaning architecturally. The meaning of the sign is determined as a function of values of significant variables or determinables exhibited by the sign. Put another way, the meaning varies systematically to parallel significant (mathematical) transformations of the sign" (Millikan 2004, 48). The correlation between an infosign and its signified is a mapping function (in the mathematical sense) from the set of the infosign's properties to the set of the signified's properties. Millikan thus calls this correlation a semantic mapping function. It yields the productivity of infosign systems—their capacity to provide new information. That is, every mapping relation between the two sets is informational and thus employable for

organisms. Furthermore, like Dretske, Millikan argues that an infosign of an affair is itself an infosign of the same affair (Millikan 2004, 53). Natural information carried by an infosign is also embedded in an infosign of the infosign. “The semantic mapping functions for such embedded signs result from applying the semantic mapping functions of succeeding signs on the route one after the other” (Millikan 2004, 54).

Infosigns carry natural information and thus cannot be false. They do not carry, but ground, non-natural information. So, where does the normative aspect of non-natural meaning come from? Millikan argues that non-natural meaning is carried by intentional signs. Different from Dretske, who appeals to individual development, Millikan takes an etiological explanation. According to this explanation, an organ has an effect on its function because it was preserved due to the effect in history, either by learning or evolution (Wright 1973). Accordingly, carrying non-natural meaning is a stabilized function of intentional signs. Just as biological functions are stabilized through dynamic interactions between sign producers and consumers, the functions of intentional signs are stabilized in the same manner. When a sign is employed by a consumer, as long as the sign corresponds to the world affair which fits the consumer’s purpose, it will continue to be employed and stabilized. Therefore, it is not the purpose of intentional signs to carry natural information. When they perform their functions by their normal mechanisms, they produce infosigns. Intentional signs are conventional in this sense. Moreover, Skyrms (2010a) formalizes the convention of intentional signs by evolutionary signaling games. Since they are designed in an etiological way through learning or evolutionary dynamics to perform certain functions, intentional signs can fail to carry the natural information they are supposed to. Therefore, the normative aspect of meaning is explained. Furthermore, Millikan argues that linguistic signs emerge from infosigns in the same way.

Neander (2017) and Shea (2018) further develop Millikan’s biosemantics extensively. Since this book is more about information while Neander (2017) just assumes a conception of information without explaining it, I will only introduce Shea (2018) here. Shea argues that to understand representational content, it is individually necessary and jointly sufficient for a theory to have two parts: The functions performed by an organism or other system, and an internal organization that capitalizes on exploitable relations—relations between internal states and the world that are useful to the system (Shea 2018, 75). Inspired by Millikan, he develops a concept of exploitable correlational information, using probability theory to characterize situations where, in a local domain, when an event happens, the probability of another event occurring changes for a univocal reason. In this case, the event carries exploitable correlational information about the other (Shea 2018, 78). For the origin of normativity of meaning, he develops the concept of task function. “An output F from a system S is a task function of S , if (a) F is a robust outcome function of S ; and (b) (i) F is a stabilized function of S ; or (ii) S has been intentionally designed to produce F ” (Shea 2018, 65). Although there are important differences between Millikan’s biosemantics and Shea’s theory, I treat them as

members of the same family of teleosemantics. Next, I will take Millikan's theory as an exemplar of teleosemantics to examine its shortcomings.

Millikan's biosemantics is powerful, but it also faces several challenges. First, according to Millikan, natural information consists of locally recurrent correlations. As the argument against Stampe and Dretske goes, anything can correlate with anything else if we change the coarse grain of the reference class. This leads to an infinite paradox: An infosign can be about anything, and therefore, it is about nothing. The embedding principle may also lead to a paradox similar to the principle of nested relations proposed by Dretske. Since correlation can be continuous, there could be an infinite embedding of infosigns about the signified affairs. Second, since an infosign can be anything, its content cannot be determined by itself. This faces the content determination problem: How can we decide the content of a sign without an interpreter? [Millikan \(2017\)](#) develops a metaphysical description of the world to provide a realistic foundation for natural information. However, it is still hard to solve these two problems. Third, Millikan argues that for an event to be an infosign of another event, the two events should be members of two corresponding recurrent families. In Millikan's terms, an infosign is a locally recurrent natural sign. Furthermore, infosigns are affordances for organisms to become intentional signs. However, organisms often take a single correlation as an intentional sign. Millikan argues that the semantic mapping function of these kinds of signs is derivative, but we still need a full explanation of how it is derived from other signs. Fourth, according to Millikan, although intentional signs are defined in cooperative producer-consumer systems, it is the consumer who determines their meaning. On the other hand, she requires that infosigns ground intentional signs, even though it is not the purpose of intentional signs to carry natural information. This implies that accidental correlations could become intentional semantic mapping functions if they meet the consumer's purpose. This suggests that the infosigns defined by Millikan are not necessary for being intentional signs, or perhaps we should not define infosigns as locally recurrent correlations for some reason. However, the consumer does not need a reason.

Although these theories understand the requirements for informational relations differently in terms of correlation, strength, reference class, etc., they share an ontological commitment: Natural information is in nature and grounds conventional meaning. Put another way, meaning, representation or intentionality must have a naturalistic foundation. This understanding faces several challenges, as I have reviewed above. This is not a fair treatment of all theories and discussions, and many important contributions have been left out. However, it is enough to help us identify the problems when trying to ground meaning, intentionality or representation in information. To provide a satisfactory account of meaning, we must solve or eliminate these problems.

A radically different understanding of meaning, which has been overlooked by the mainstream of naturalistic philosophy, is Peirce's theory of signs, or semiotics.³ A reason philosophers have refused to consider this approach is that it understands every sign as having an interpretant—something that interprets it and translates it into other signs. In other words, interpretation is necessary for

something to be a sign. This seems subjective and thus appears incompatible with naturalism. Nevertheless, I believe that Peircean semiotics may provide a novel way to understand meaning with its modern naturalization, biosemiotics, without facing the challenges mentioned above. In this book, I defend the idea that meaning grounds in information, but within a biosemiotics framework, especially the emergent dynamics theory and biosemiotics proposed by [Deacon \(1997; 2012a\)](#). The accounts provided in the book comprehensively explain the conception of information, offering a naturalistic foundation for understanding intentionality, meaning, reference or representation. By naturalizing the conception of interpretation, this naturalistic theory can solve or eliminate the problems faced by past naturalistic theories. These problems include the sense and reference problem, the content determination problem, the organism use problem (how a correlation can serve as a sign relation for an organism's usage), the misrepresentation/misinformation problem (or the normativity problem, how the normativity of representation arises from nature), the reference class problem, the strength problem, the singular relation problem, the language emergence problem, the fiction problem and more. In the next section, I will present the structure of the book.

1.5 Overview

[Chapter 1](#), “Grounding meaning in information”, introduces the core problems that the book aims to address: providing a naturalistic foundation for intentionality by explaining the conception of information. The chapter critically reviews various attempts to characterize informational content and naturalistically explain intentionality. Finally, it presents the structure of the book.

[Chapter 2](#), “Structural, referential and normative information,” provides a comprehensive conceptual analysis of information. It begins with a folk notion that information is a tripartite phenomenon: Information is something carried by signals about something for some use. This suggests that information has three main aspects: Structural, referential and normative. Individually necessary and jointly sufficient conditions for defining each aspect of information are articulated within the analysis. Formal theories relating to each aspect are also considered. Finally, the analysis reveals that structural, referential and normative aspects of information are hierarchically nested and that the normative depends on the referential, which in turn depends on the structural. This lays a conceptual ground for a naturalistic account of meaning and representation.

[Chapter 3](#), “Information is physical (negatively),” develops a naturalistic account of structural information. This chapter first argues that understanding information as the communication of form revives the etymological meaning of “information.” Therefore, the key to understanding information lies in how we understand form. Deacon proposes that form is an intrinsically absent aspect of something present. This view helps address a longstanding puzzle in information studies: The superficial similarity between the formulas for entropy in Shannon's mathematical theory of communication and in thermodynamics. It reveals the

physicality of information and lays a physical foundation for understanding reference and intentionality. However, unlike fundamentalist views of information, I argue that the physical ground of information is not equivalent to information itself, or we may encounter issues such as misinformation and content determination.

Chapter 4, “Interpretation,” explains the origin of normativity. It begins with the argument that normativity is implied in both information and meaning. The chapter then critically reviews the leading approach to explaining the normativity of meaning in naturalized semantics, namely teleosemantics. It argues that the primary advantage of teleosemantics is that it does not require teleological notions to explain normativity, relying instead on the selected effect account of function. However, it faces the problem that what the selected effect account explains are the effects of a biological trait or item produced in the past (history), whereas what it produces are effects in the present. As an alternative, this chapter proposes a naturalized Peircean semiotic account in which interpretation plays a central role in explaining the normativity of meaning. It supports Deacon’s concept of autogenesis as a plausible model for a primary interpretive system. Finally, the concept of operational interpretation is introduced to understand interpretive processes at the unicellular level.

Chapter 5, “Reference,” presents a semiotic analysis of reference. It first critiques the map metaphor in traditional philosophy of language and philosophy of mind. The chapter then argues that, according to Peircean semiotics, reference is the result of hierarchically organized interpretive competence. Symbolic interpretation, it suggests, asymmetrically depends on indexical, which in turn depends on iconic interpretation. The chapter argues that a symbol differs from an icon and an index because it has no direct correlation with what it represents and does not lose its referential power even when it loses its direct correlation with the represented object. The chapter further claims that symbolic reference retains this feature because it is ungrounded from the direct correlations with its referent, instead discovering the indexical relationships between symbols. This feature explains why abstract, generalized and fictional descriptions are possible. Finally, the chapter argues that for a symbol to represent an object in the world, it must be regrounded by correlating indexical relationships that are directly correlated with the object in the world. This explains the functions of different types of words.

Chapter 6, “Convention,” argues that Lewis-Skyrms’s evolutionary sender-receiver signaling game theory is insufficient to explain the conventionality of meaning, proposing the mutual misunderstanding argument as a counterpoint. The chapter develops a semiotic account of convention, first reformulating the evolutionary signaling game theory, then presenting the mutual misunderstanding argument through a thought experiment, *the Magical Fight*. It further argues that the conventionality of meaning has dual aspects: Both the sign vehicle and the referential relationship of language are conventional. The chapter asserts that what is explained by evolutionary signaling game theory is the convention of the sign vehicle, and with a semiotic understanding of convention, we can better explain scenarios of mutual misunderstanding.

Notes

- 1 This is an oversimplification of Floridi's measure of the quantity of semantic information. The reformulation here tries to capture the main points rather than an accurate reconstruction of the theory. The original measure is more complex. People who have interest in it may read [Floridi \(2004\)](#).
- 2 Reduction in uncertainty is one way to understand information. Information can also be understood as reduction in doubt, divergence and expected inaccuracy. How to characterize these understanding formally is open to interpretation ([Roche and Shogenji 2018](#)). I focus on Shannon's understanding as it is the most common one and other understandings do not affect my arguments substantially.
- 3 This claim is not fair enough. [Millikan \(1984, 85; 2017, 144\)](#) does have some empathy in Peirce's theory of signs. Her theory of signs is inspired by Peirce in some sense. Nevertheless, she rejects involving interpretation in natural information.

2 A conceptual analysis of information

2.1 Introduction

Compared to concepts such as “meaning,” “reference,” “representation,” “value” or “significance,” most scientists would prefer the concept of “information.” This is because those concepts are philosopher’s playthings, replete with philosophical confusions. Thanks to Claude Shannon, Andrey Kolmogorov and others, information can be quantitatively defined with mathematical precision and represented in formal theories. This development was essential for the development of both computer science and communication engineering (Kolmogorov 1963; 1968). Building on this success, scientists believe that the concept of information more generally should be scientifically valid when applied to physics, biology, cognitive science and social science. Unfortunately, however, confusion and misunderstandings still arise. Over time, such usage widened and scientific knowledge went far beyond the parameters of formal theories. That is, information became confused or conflated with other notions, including the “message,” “data,” “computing,” “code” and “meaning” (Capurro and Hjøland 2005; Brenner 2014).

Instead of criticizing their inappropriateness and their careless use, I believe that there is a reasonable motivation to impel such usages of the concept of information. That is, those usages do involve some aspects that are essential to information but not covered by those formal theories. Therefore, a crucial task of information studies today is to provide an explanation of information that can help us understand the basic aspects of information and the relationships between them. In order to explain the concept of information, two kinds of work need to be done. First, there is conceptual analysis, “.....in a paradigm case, an analysis embodies a definition; it specifies a set of conditions that are individually necessary and jointly sufficient for the application of the concept. For proponents of traditional conceptual analysis, the analysis of a concept is successful to the extent that the proposed definition matches people’s common sense about particular cases, including hypothetical cases that figure in crucial thought experiments” (Fallis 2015; Margolis and Laurence 2019). Accordingly, the strategy I use in the chapter has the following steps: (1) Proposing an imagined scenario of information that is common in everyday life; (2) deriving a folk notion of information from the scenario; (3) figuring out the set of conditions that are individually necessary and jointly sufficient for each aspect

of information through conceptual analysis; (4) and discussing the implications of the definitions for formal theories for each aspect of information. Why conceptual analysis is necessary for explaining information? While the concept of information is widely used in science, “[i]t is,, employed somewhat differently in each, to the extent that the aspects of the concept that are most relevant to each may be almost entirely non-overlapping” (Deacon 2007). This points to the need to articulate the meaning of information both clearly and conceptually. This is what conceptual analysis can provide. With conceptual analysis as the starting point, we can engage with the second kind of work: Theory construction. Unlike the conceptual analysis, being analytic that specifies the set of conditions defining each aspect of information, theory construction is synthetic, which aims to provide explanations about the mechanisms underlying the processes of information. Nevertheless, this is not the place to fully unfold the theory and the chapter focuses on the former.

The chapter argues that information has three unreduced aspects, structural, referential and normative and specifies the individually necessary but jointly sufficient conditions for defining each aspect through the conceptual analysis. The analysis also shows, as a result, that those three aspects are hierarchically dependent that normative information depends on the referential which depends on the structural. The thesis argues against pan-informationalism, which claims that information is monistic and everywhere in the cosmos. The structure of the chapter is arranged as such, in Section 2.2, I begin with an imagined scenario to illustrate a folk idea that information is a tripartite phenomenon: Information is something carried by *signal about* something *for* some use. In this conception, information has three aspects: Structural, referential and normative. The conceptual analyses of structural, referential and normative information are made in Sections 2.3, 2.4 and 2.5. At last, the nested hierarchy of information that normative information depends on referential information, which depends on structural information, is revealed. Furthermore, possible objections to specific claims of each aspect are dealt with in each section.

2.2 Information as a tripartite phenomenon

There is a folk notion that information is something carried by *signals about* something *for* some use.¹ How to understand this everyday notion of information? A case provided by Bar-Hillel gives us a good start illustrating the basic aspects of information captured by the common sense:

A writes on a sheet of paper “I love you” and wishes that *B*, 3000 miles away, should become aware of the full content of this message, with little delay and at a low cost. *A* will be dissatisfied if he learns either that his message has been scrambled up, whether into something incomprehensible like ‘K bogl pou’ or into something comprehensible like ‘a longbow’ but with entirely different meaning, or that an undistorted replica has been delivered a day late (and of course, even more so if a distorted message is delivered too late).

(Bar-Hillel 1955)

The case describes a scenario of communication in which information is transmitted. I call the scenario $A \heartsuit$ (love) B . In $A \heartsuit B$, there is an informer or source, A , two converters, the telegraph machines encoding and recoding the message, the channel which transmits physical signals encoded by a converter and an informee, B . The whole scenario constitutes a communication system, whereby what is transmitted from A to B is information. In other words, information flows from A to B through the chosen system of communication. Furthermore, communication and information are distinct. Communication is the transmission of information from source to receiver concerning about the origin and end of information. Nevertheless, if we want to understand information, it should be discussed in the context of communication.

What the setup describes is a linguistic communication between two human beings. Human linguistic communication is the most sophisticated information process that we know. Synchronically, information processes in the living world are much richer and more diverse than linguistic communication. Apart from the linguistic form of information processes, there are various extant signaling phenomena, ranging from chemical signals, employed by bacteria and plants, to the signals that animals employ to communicate. In addition to the information processes at the individual level, there are also others at sub-personal levels, including genetic information, hormones, etc. Furthermore, the relationship between the information and its linguistic carriers is conventional. There is also natural information, for example, smoke carries the information of fire, dark cloud conveys the information of raining and tree rings embody the information of the age of trees. Grice (1957) was the first to distinguish natural from non-natural meaning. Grice was followed by Millikan (2004) in defining information as both intentional and natural signs. Do all these events have something in common when treated as being informational? If not, why should we seek to all of them in a concept of information? What are the differences and connections between linguistic and other forms of information processes? There is no doubt that these are crucial questions of information. However, the conceptual analysis of information is far from being enough to answer these questions. Instead, it requires a comprehensive theory of information. Consequently, these wider questions are out of the scope of this limited paper. On the other hand, the purpose of this chapter is not irrelevant to the wider goal, since the construction of a theory of information should take conceptual analysis as a premise. In other words, making a blueprint is a prerequisite, a first step, in building a concrete edifice. The conceptual analysis is the blueprint of the theory construction.

In theories of evolution, it is widely believed that the linguistic capacity of *Homo Sapiens* evolved from the primary information processing capacities of primitive organisms (Deacon 1997). Yet, by taking this scenario as an exemplar of information processes, the whole evolutionary history of information is skipped. As a result, some may argue that this chapter forgoes the possibility of information unification via an unfolding process of increasingly sophisticated information processes. Yet, this chapter agrees that the theory construction of information would considerably benefit from illuminating the evolutionary history of information.

However, I believe that to begin with a daily scenario is a much easier way to think of a complex phenomenon.

When we say that we collect, store, transmit and retrieve information, it implies an obvious but widely accepted notion: Information is an objective commodity whose generation, transmission and reception are independent of any pre-existing interpretive processes and transferred from senders to receivers through channels in communication systems (Dretske 1981, vii). It is the message “I love you” that is transmitted from *A* to *B* via the telegraph in the above scenario. The message is an objective commodity that what the symbols represent, the content of the message and the physical signals in which the message are coded and transmitted are independent of any preexisting interpretive processes.

Given that the message is transmitted without distortion, is the message information? Well, it is hard to say. There are several potential scenarios whereby no information is transmitted even though there is a message. First, if *A* actually does not love *B* and is just making a joke, then *B* receives not information but misinformation. It will be argued later that misinformation is not information. Second, if *A* does love *B* but *B* already knows it, then the message is not information as *B* has already known it, given other things equal. Third, if the message accidentally is sent to another person, *D*, then the information transmitted is “*A* loves *D*,” given *D* believes that the message is supposed to send to her/him. So, it is clear that the *information* transmitted from *A* to *B* is not necessarily identical to the meaning contained in the message. In these cases, information is confused with the meaning of or the words representing the meaning of the message while information is not identical with them.² Then, what is information? *A* ♥ *B* scenario helps.

In *A* ♥ *B*, *A* will be dissatisfied with three hierarchically nested situations: (i) Distorting his or her message into incomprehensible signals; (ii) the message contains comprehensible signals but which a different meaning for *B* than intended by *A*; (iii) or delivering the message too late to be useful. Say the three situations are hierarchically nested means that, when a message satisfies its expected purpose (usefulness), it presupposes the message conveys the correct meaning, which presupposes that the symbols representing the meaning are not distorted. These three situations actually correspond to three basic nested aspects of information. These three aspects are *signals*, *aboutness* and *forness*. The usefulness of information presupposes aboutness, which presupposes signals. In turn, the usefulness hierarchically depends on aboutness which hierarchically depends on signals. The hierarchically nested dependence of these three aspects will be further articulated in the conceptual analysis of each aspect. It is consistent with our everyday usage of information. We have a variety of terms to talk about those three aspects: Syntactic, semantic and pragmatic information; measurement, meaning and usefulness of information; quantity, content and value of information, and so on.

“It is hardly to be expected that a single concept of information would satisfactorily account for the numerous possible applications of this general field” (Shannon 1993, 130). Shannon can be regarded as the originator of the widely accepted idea that information has three aspects. Weaver’s classification is also well-known.

He formulates his classification in terms of communication. If a piece of information is to be transmitted successfully via communication, then three levels of problems must be solved: Technical, semantic and effectiveness. The technical problem concerns the accuracy of the transmission of the symbols of communication. The semantic problem relates to “the identity, or satisfactorily close approximation, in the interpretation of meaning by the receiver, as compared with the intended meaning of the sender” (Shannon and Weaver 1964, 4). The effectiveness problem focuses on “the success with which the meaning conveyed to the receiver leads to the desired conduct on his part” (Shannon and Weaver 1964, 5).

Contemporary researchers have also proposed their conceptual classifications for specific purposes since the 1950s. I’ll just review a few instances among many here. Bates’ classic works (2005; 2006) classify information into three forms: Information 1 (the pattern of organization of matter and energy), information 2 (pattern with meaning) and knowledge (mental states of a mind as the result of being informed by the pattern with meaning). Floridi (2011) argues that information can be viewed from three perspectives: “information as reality (e.g. as patterns of physical signals, which are neither true nor false); information about reality (semantic information, alethically qualifiable); and information for reality (instructions, like genetic information, algorithms, orders, or recipes).” Deacon (2017) uses physical, referential and normative information to refer to these three perspectives. Harms (2006) argues that there are three crucial concepts of information: Statistical, semantic and physical. Similarly, Gregersen (2010) formulates these three as counting, meaning and shaping information. It is necessary to note that, when listing the accounts of information above, it did not imply that the tripartite division mapped to each other accurately. Nevertheless, it is obvious that they have some similarities. The question might be asked: Do we need another attempt to clarify the concept of information? A lack of an adequate conceptual analysis of information is a major part of the concern. The existing conceptualizations of information are lacking in some way with respect to conceptual analysis.

For examples, in Bates’ (2005; 2006) and Linski’s (2010) formulations, knowledge as a mental state is defined as the result of one’s mind being informed. Surely, knowledge is deeply involved in the normative aspect of information, as we will see in Section 2.5. However, it is not the only result of one being informed. There are also changes of one’s actions. Floridi’s insightful classification and his presentation of the map of information covering structural, referential and normative information inspired my research a lot. The classification even implies the hierarchy of those three aspects in the map of information he presents (Floridi 2010, 18). Nevertheless, more comprehensive and explicit analysis, I believe, is still needed. A problem of Harms’ theory is that statistics is a property of physical (structural) information, not in parallel with it. This will be argued in Section 2.2. Nevertheless, the analyses presented in the chapter benefit a lot from those past works and even in which many ideas are directly cited from them. What the chapter originally contributed to information studies is that it tries to integrate those ideas through the rigorous method of conceptual analysis and thus to specify the individually necessary and jointly sufficient conditions for defining each aspect of information.

In the history of information studies, the most enduring framing of the three aspects of information is syntactic, semantic and pragmatic information though people also use other terms as being shown above. These terms are borrowed from linguistics originated from Charles Morris. This is a language-centric way to frame the classification (Deacon 2017). However, this linguistic framing of information theory has several shortcomings. First, it is obvious that terms are at the center of debates in contemporary philosophy. They are heavily theoretically loaded and thus not a good starting point to begin to understand such concepts. Second, it is widely believed that information is more basic than language. Linguistic symbols are just one kind of sign that conveying information among many. Language is a special derived case of information, nor a generic one. Furthermore, applying linguistic glossaries—like syntax, semantics and pragmatics—to information studies may mislead us into confusing linguistic symbols conveying conventional meaning with natural signs carrying natural meaning (Grice 1957; Millikan 2004). This would obscure critical semiotic distinctions between them and overlook a possible continuum involving them. This chapter adopts Terrence Deacon’s terms, namely, structural, referential and normative information, to represent these three aspects of information. Next, let us turn to the three situations that dissatisfy *A*, in order to illustrate these three aspects of information.

2.3 Structural information

In $A \heartsuit B$, *A* is not satisfied with (i) because his message is distorted into an incomprehensible symbol sequence. This concerns the pattern of symbol sequences or signals carrying information. It is obvious that the pattern of signals can be multiply realized by different physical media. In the case of $A \heartsuit B$, the pattern of “I love you” is realized by ink on the sheet of paper written by *A*, or by electric signals during the transmission by telegraph. Furthermore, it is impossible to imagine that there is a pattern without it being manifested by physical medium—unless you believe Plato’s account that *eidos* is real. That is to say, patterns are realized but not determined by the physical. We call this pattern of signals, structural information.

In order to provide a definition of structural information that necessarily and sufficiently explains it, we should first specify the necessary elements of it, check the possible relationships between them and then argue which are necessary for it. Jointly, the elements and relationships between them are sufficient for structural information because all the necessary elements and relationships between them are included. Moreover, the conceptual analyses of referential and normative information also follow the strategy. In the case of structural information, as we can see, *ceteris paribus*, there are two necessary elements: Pattern and its physical realizer(s). Then, the definition of structural information should answer the following questions: What is a pattern? What is the relationship between a pattern and its physical realizer(s)? Does a pattern identify with its physical realizer(s)? If not, what is it?

With respect to the questions, we call an aspect of information structural information (*Str*) if and only if it has the following properties: (a) it is the pattern or

difference manifested by the physical medium; (b) it is realized by the physical but, nevertheless, the physical is neither sufficient nor necessary for it; (c) it is concrete abstraction.

(a) states that *Str* is the formal aspect of information. In Floridi's terms, we can understand *Str* as difference *de re* metaphysically, "i.e. mind-independent, concrete points of lack of uniformity (Floridi 2011, 356). Floridi (2010, 17) defines data as the distinction between two uninterpreted variables. The lack of uniformity, or difference, comes from symmetry breaking (Collier 1996). Since *Str* is the lack of uniformity in a physical medium, it is realized by the medium's physical properties.

In this scenario, *Str* refers to the linguistic symbol sequence, "I love you," Morse Code, and electronic signals. Those are arbitrary products intentionally produced to serve certain functions. However, we should not be misled into thinking that *Str* is limited to arbitrary sign vehicles. Where there is symmetry breaking or change of state, there is *Str*. Examples of *Str* include agitated states of fundamental particles, a gamma-ray burst of a stellar system far away from Earth, smoke, dark cloud, etc., are all *Str*.

(b) states the relationship between *Str* and its physical medium. Informational relationships surely depend on underlying physical processes but do not identify with them.³ There is no problem that *Str* is realized by physical properties. For example, the same symbol sequence, "I love you," is realized by physical properties, not matter of *A*'s voice, ink on a paper, or electrical signals. Thus, *Str* depends on physical processes. It is also clear that the sequence is multiply-realized by different physical properties. Thus, *Str* is not identified with physical processes. Furthermore, a physical relationship is neither necessary nor sufficient for an informational relationship.

A physical relationship is not necessary for relaying informational. Two events, E_1 and E_2 , have a physical relation when they materially or energetically connect with each other. In terms of *Str*, E_1 and E_2 have informational relation when the *Str*, as known as "form" or "pattern" of E_1 , can be detected from E_2 . Physical events, of course, can be informative. A gamma-ray burst detected by a Fermi Gamma-ray Space Telescope conveys the *Str* of a stellar system which is billions of light-years away from us because the former is physically caused by the latter. The ring of a doorbell tells of the arrival of a visitor outside of the door because the ring is physically activated by the visitor. However, *Str* is not determined by physical relations. It will be argued later that, according to Shannon's theory, the quantity of the generated *Str* of a physical event is not determined by the events itself. That is to say, the quantity of *Str* generated by the event is not determined by the physically present event but by physically absent ones. Some may argue against this that only when the event physically happens can there be *Str* even though the quantity of *Str* is not determined by the physical presence of the event. Therefore, it seems that a physical relation is necessary for *Str*. However, even absent physical events can be informative. "No news is good news!" In addition, imagine a boiler has a reliable alarm. If the pressure inside the boiler increases beyond the safety threshold, the alarm will be triggered via both sound and light. But if the alarm is not triggered, it indicates that the pressure of the boiler is below the safety threshold.

A physical relationship is not sufficient for informational relations. Although any causal sequence of physical events can be *Str*, the sequence itself cannot determine which *Str* matters. For example, in the case of a doorbell ringing, the ring is the signal that conveys the information of the arrival of a visitor. At the same time, the process is a physical causal process: The sound heard by the person inside the dwelling is caused by the vibration of the ring consequent to the electric current, the result of the interaction between two pieces of sheet metal located in the doorbell, which is itself triggered by pressing the doorbell by the visitor. Obviously, there are many physical events which happen in the entire process and the differences or forms realized in any one of them are *Strs*. However, we believe that the ring conveys the *Str* of the arrival of the visitor rather than other *Strs*. It means that a detailed explanation of the causal chain of the process cannot explain what *Str* is conveyed in the chain of events.⁴ This implies that a physical relationship is not enough to be an informational relationship. This will be further argued in [Sections 2.4 and 2.5](#).

Since any difference or constraint is *Str*, *Str* is ubiquitous in the universe. There are scholars who argue that information is *Str* ([Zuse 1967](#); [Wheeler 1989](#); [Stonier 1997](#); [Schmidhuber 1997](#); [Dodig-Crnkovic 2011](#)). The view is called pan-informationalism or digitalism. The argument of this chapter is that such a view is untenable. This is because, first, if the information identifies with *Str*, then we cannot distinguish information from physical events, which always have certain structure. Second, information has the property of intentionality, that is, it is about *something*. The signal sequence, aka *Str*, conveying content is not equal to the conveyed content ([Carnap & Bar-Hillel 1952b](#); [Bar-Hillel 1955](#)). Even if a physical event provides *Str*, it tells us nothing about either the referential relationship between it or its reference and significance. Therefore, information is more than *Str*. *Str* is potentially informational but is not itself information. Further, if information is just form, pattern, difference, organization, structure, etc., then, why do we still talk about information? In that case, the concept of information would be redundant. One reason that we need the concept of information is that it is different from physical events.

If *Str* does not identify with its physical medium, where is its place in the physical world? (c) explains what *Str* is. A long-lasting metaphysical question relating to *Str* is that where is the place of form or Platonic *eidos* in the physical world? It seems that forms, patterns, or structures are something that are added to the physical world. We have only two ways to answer the question: First, following Plato, *Str* is something abstract but real that is essentially different from the physical. Second, *Str* is something interpreted by observers, it is not real in the sense of the physical and only exists in observers' minds. Since the first option has long been rejected, the second one seems to be the only choice. That is to say, differences, patterns, forms, structures, organizations, etc., are observer-dependent, not real ([Deacon 2012a](#), 187–189). However, the option leads to a dilemma: On the one hand, it is anti-intuitive that we believe that *Str* is as real as triangles, rectangles and circles in common sense; on the other hand, if it is right, it has to solve a harder problem, how could mind produces *Str*? The problem implies a regress: “[T]o

attribute physical regularity [aka. *Str*] to some perceived or measured phenomenon presumes a prior mental regularity or habit with respect to which the physical regularity is assessed” (Deacon 2012a, 189). Therefore, the option only postpones the problem rather than solves it.

Deacon uses “constraint” to try to explain away the dilemma. Technically, constraint is a term used to describe some reduction of degrees of freedom of change or a restriction on the variation of properties that is less than what is possible. Using constraint to define *Str* in a negative way saves *Str* both the feature of being abstractly general and of being concretely real.

The general logic is as follows: if not all possible states are realized, variety in the ways things can differ is reduced. Difference is the opposite of similarity. So, for a finite constellation of events or objects, any reduction of difference is an increase in similarity. Similarity understood in this negative sense — as simply fewer total differences — can be defined irrespective of any form or model and without even specifying which differences are reduced.

(Deacon 2012a, 190)

In short, constraint or *Str* is the elimination of certain specific features that could have been presented. *Str* is abstract general because this formulation preserves the feature that *Str* is the similarity of different particular objects or events. *Str* is real because it is the result of the elimination of particular features that could have been presented, which is the result of certain specific particular processes or events. This will be fully argued in chapter 3.

Since *Strs* are the results of particular processes, they must have an extension implying that it is measurable. *Str* is measurable in two senses: intrinsically and extrinsically. In the case above, *Str* is the symbol sequence “I love you.” It is composed of ten tokens including eight letters and two blank spaces. Since the aim of the telegraph company is to gain as much interest as possible in sending the message, then the company has a vested interest in providing the service as cheaply as possible. To do that, the company needs to find the shortest way to encode or describe the sequence, or, to measure how much *Str* the message contains in terms of the number of bits required to describe it. In other words, the complexity of the sequence should be known in order to transmit it most efficiently (Li and Vitányi 2008, 101). The way to measure the quantity or the complexity of *Str* is the Kolmogorov complexity or algorithm information independently established by Kolmogorov (1963; 1968), Solomonoff (1964a; 1964b) and Chaitin (1987). We can see that the complexity of a signal sequence is determined by itself and we do not need to consider other things than the sequence itself to measure its *Str*. Therefore, we say *Str* is intrinsically measurable in this sense. This is different from Shannon’s way of measuring the quantity of *Str*, which is extrinsic.

In Shannon’s mathematical theory of communication, information is defined as the reduction of uncertainty (Shannon 1948). The degree of uncertainty is determined by how many alternatives there are. The more alternatives, the more uncertainty. Thus, the more alternatives are reduced, the more information is carried by the sent signal. That is to say, the quantity of *Str* carried by a signal is not

determined by the signal itself but by alternative signals that might have been sent. Suppose there are two equally possible states of a source, S_1 and S_2 , and two available signals, M_1 and M_2 . M_1 represents S_1 and M_2 represents the other. If S_1 actually happens, then the quantity of information carried by M_1 is 1 bit according to Shannon's formula. Given S_1 and S_2 , M_1 and M_2 unchanged and suppose that there are two extra possible states and two signals, then M_1 carries 2 bits of information rather than 1 bit. In these two situations, S_1 and M_1 do not change, but the quantity of information generated and transmitted changes. Therefore, the quantity of information measured by Shannon's theory is determined extrinsically rather than intrinsically.

2.4 Referential information

In $A \heartsuit B$, (ii) is not satisfactory for A because the delivered message has a different content from the original, although the transmitted Str is meaningful and thus comprehensible. The symbol sequence "I love you" as a sentence has the same meaning with the symbol sequence "Je t'aime" although their patterns are different. Thus, the referential content of a message does not identify with the pattern, namely Str . In other words, there is no intrinsic relation between the content of a message and the pattern of a signal's medium. B will know A 's true belief about the love between them through the message and thus reduce her uncertainty on A 's feeling on her. This is referential information.

In the case of referential information, *ceteris paribus*, there are three necessary elements: Str , the referential relationship and the object or event the Str refers to. Then, the definition of referential information should answer the following questions: What is a referential relationship? Is a referential relationship determined by the intrinsic properties of Str ? Is a Str dedicated to referring to an object(s) or event(s) intrinsically? Or is a referential relationship determined by the referent?

With respect to the questions, we call an aspect of information referential information (Ref) if and only if it has the following properties: (a) it is the relation of a signal (Str) to the object it refers to; (b) it depends on Str that the latter is its vehicle but not determined by it; (c) it is intrinsically alethically neutral.

(a) states that Ref is the aboutness aspect of information, that is, the referential relationship between information and the things it refers to. Roughly speaking, when the probability of the occurrence of a signal given the referent the signal (Str) refers to is different from the probability of the occurrence of the signal *per se*, we say that there is a referential association between the signal and the referent (Shea 2018, 76). The referential relationship is not intrinsic between the signals and things the signals are about. For example, the Ref carried by the message "Jinping Xi is the president of the People's Republic of China in 2020" has no direct physical relation to the event that "Jinping Xi is the president of the People's Republic of China in 2020." Furthermore, the message given above can be expressed by different languages (symbols) in different physical forms (e.g., sound, ink, electric screen, etc.), it means that Ref is multiply realizable. As a result, Ref has no intrinsic

relation to *Str*. This is what (b) claims. Think of (b) in the other round, the definition of *Ref* presupposes *Str*. Therefore, we can say *Ref* hierarchically depends on *Str*.

It is easy to confuse *Ref* with meaning. Despite their superficial similarity, *Ref* is different from meaning. First, meaning is a polysemantic concept, as the philosophy of language shows. When we say, for instance, “What is the meaning you want to express?,” we are pointing to the semantic content of a sentence. “What is the meaning of life?” refers to the significance or value of life. The connotation of *Ref* is actually much clearer although not entirely free from confusion. Second, while the semantic content may be close to *Ref*, it is a linguistic-centered concept, whereas *Ref* is more fundamental and is not limited to linguistic phenomena, as already argued above.

Third, the *Ref* conveyed by a symbol may not identify with its conventional meaning.⁵ On the one hand, the *Ref* conveyed by a sign may exceed the conventional meaning of the sign. For example, when I see a sentence on a paper, I may get the information of its original source once I read the sentence in its original text and remember it. “.....[H]ence the information carried by that signal depends in part on what one already knows about the alternative possibilities” (Dretske 1981, 43). This is what exactly Shannon (1948) tells us about in his theory of information. On the other hand, natural signs like the footprints of animals on earth, a tree’s growth rings, dark clouds, someone’s facial expression and so on, have no conventional meanings. Yet, they surely carry *Ref*. Furthermore, symbols which have meaning may not convey any *Ref* at all. In $A \heartsuit B$, the message “I love you” sent to *B* by *A* is information for *B* only when *B* does not know that *A* loves him or her. If *B* already knows that *A* loves him or her, then the message is redundant not informational for *B*, although there is no change of meaning. Moreover, if I speak to my wife that “I am busy writing my book,” while I am actually playing a game on my mobile, then the words convey no information to my wife because I am not telling the truth. In other words, the meaning represented by my words does not correctly correspond to the facts. The scenario implies that *Ref* is necessarily true. This is so-called the veridicality thesis.

I agree with the veridicality thesis. However, misrepresentation is not unusual. A spy deceives enemy personnel by intentionally spreading false information (disinformation). An unaware guard mistakes a stranger for a leader (misinformation).⁶ Intuitively, we do not think that there is any information conveyed in these two situations. Why does information have to be true? According to a formulation of Shannon’s communication theory, information is the reduction of one’s uncertainty about a subject by eliminating alternative possibilities (Berger and Calabrese 1975). False information cannot eliminate alternative possibilities and reduce one’s uncertainty. Suppose *B* loves *A* but does not know whether *A* loves him or her or not before the scenario takes place. In other words, *B* is uncertain about *A*’s feelings. After receiving *A*’s message, *B*’s uncertainty is reduced on the condition that the message corresponds to the fact that *A* loves *B*. If *A*, in fact, does not love *B* and sends the message to play with the recipient’s feelings, then *B* may be misled to personally believe that the uncertainty is reduced but not actually. In the scenario, it is disinformation. According to the misleading nature

of disinformation, Fallis names disinformation as misleading information (Fallis 2015). As a result, veridicality is an intrinsic property of information.⁷ It implies that neither misinformation and disinformation are kinds of information. Then, why do we still call them *mis*-information and *dis*-information? Dretske (2007) argues that “[t]his is a pretty heavy-handed treatment of ordinary language.” That is to say, misinformation and disinformation are not information—just like decoy ducks are not real ducks.

Superficially, the discussion of the veridicality thesis is inconsistent with (c), which asserts that *Ref* is intrinsically alethically neutral. Actually, it is not. Veridicality is a property of information but not of *Ref*. What *Ref* characterizes is just the referential association between signals (*Str*) and their referents. However, *Ref* cannot determine the relation itself because anything associates with anything in one way or another. Think of the door ring scenario proposed in the last section. The sound of the ring can be about any event in the causal chain. Therefore, to discuss the truth value of *Ref* by itself is empty. Why the ring sound is dedicated to convey the information of the arrival of visitor? Because it is designed to be so by its designer’s intention. Therefore, the problem of referent-determination of a signal is about the normative aspect of information. It is a topic discussed in the next section.

Many people believe that *Ref* is measurable and seek to construct ways to define the quantity of *Ref* (see, for examples, Bar-Hillel & Carnap 1953a; 1963b; Floridi 2004; Skyrms 2010a; 2010b). Other scholars disagree with this interpretation. For example, Dretske (1981) argues that, given the receiver already knows about the possibility of the source, only when the conditional possibility of *s* being *F* is 1, can we say that a signal carries the information *s* is *F*. If the sent and the received messages are different, even slightly, then qualitatively they are two different messages. However, the interpretation does not hold for all possibilities as it is counter-intuitive (Scarantino 2015; Stegmann 2015). For example, intuitively, the message “Luciano Floridi is a male Italian philosopher of information” carries more *Ref* than the message “Luciano Floridi is a philosopher.” But in what sense? According to Bar-Hillel & Carnap (1953b), the possibility of occurrence of the former is less than that of the latter. In contrast, because Bar-Hillel and Carnap’s approach leads to a paradox that an analytic proposition carries no *Ref* and a proposition of contradiction carries infinite *Ref*,⁸ Floridi (2004) would argue that the degree of the truth brought by the former is larger than the latter. Skyrms (2010b) would claim that the former has more *Ref* than that of the latter because the former has more vectors. The argument of this chapter is in contrast a pluralist position regarding how to characterize the quantity of *Ref*. So long as a theory of the measurement of *Ref* fulfills its designed purpose, then it is acceptable. Nevertheless, it is not clear that the claim that *Ref* is measurable conflicts with Dretske’s assertion. That two messages with minor differences are qualitatively two different messages does not mean that they are quantitatively incomparable. So long as one provides an acceptable standard of understanding, then they are comparably measurable. Qualitative differences matter in normative information, which will be examined in the next section.

2.5 Normative information

In $A \heartsuit B$, A is not satisfied with (iii) because the delayed message cannot fulfill his purpose. It cannot make the difference he or she wants, in Bateson's terms (1972). Suppose B is getting engaged to another person next day. This is despite the fact that she loves A but cannot wait any longer, because she does not know A 's mind and their family is pressing for a decision. So, A has to let B know that he loves her before the engagement, or nothing can be resolved. It follows from this that information is sensitive to time, although the delivery of words and their content does not change. More generally, information is contextually sensitive. For example, the delivered message has a different value for A compared to the telegrapher who sends it. Why A can use the symbol sequence "I love you" to express the thought to B is because the symbol sequence has the same meaning for both A and B . Put differently, the symbol sequence "I love you" is supposed to serve the function that one expresses love to the other in a two-person communication. The same is the case with *Ref*. That is, A does not care about how the message is symbolically (*Str*) delivered, wishing primarily to understand that the message was delivered correctly and on time. We call this aspect of information, normative information.

In the case of normative information, *ceteris paribus*, the necessary elements of normative information are *Ref* carried by *Str*, the effect caused by *Ref* and the value for the user of the information. Then, the definition of normative information should answer the following questions, why a signal is dedicated to be about a certain referent? What is the relationship between *Ref* and the effect it causes? What is the relationship between *Ref*, the effect *Ref* causes and the value for information users?

With respect to the questions, we can call an aspect of information normative information (*Nor*) if and only if it has the following properties: (a) it is stabilizedly useful for information users; (b) it depends on but is not determined by *Ref*, in turn, it determines the specific referential association between a *Str* and its referents, and thus determines *Ref* of the *Str*; (c) it is sensitive to contextual factors.

Nor is concerned with information's usefulness aspect. An intuitive way to describe it is that *Nor* is a difference-maker (Scarantino 2015). As Bateson (1972, 453) claims correctly, it is "a difference which makes a difference." Bateson takes advantage of the ambiguity between two meanings of "make a difference," that is, "to matter" and "to cause to change" (Deacon 2012a, 332). *Nor* is always for some end, relating intrinsically to usefulness. Only when a message can change the epistemic or action state of a receiver can we refer to it as *Nor*. For example, the message "a whale is a mammal" would be new information to people who lived in China 2000 years ago. But it is not (new) information for most people today because it is well-known and thus redundant losing its novelty. That is, the information does not change their epistemic state. It leads to (b). The value of information, aka *Nor*, is contextually sensitive with respect to time, individual receivers and other factors. For example, "Terrence Deacon is a faculty member of the Department of Anthropology, University of California, Berkeley" is not *Nor* for me, but it is for my mother. This is what (c) asserts. It will be further explained later.

(a) not only states that *Nor* is useful for users but also stabilizedly useful. This is called proper function or stabilizing function of a message (*Str* + *Ref*) by Millikan (1984a; 2005). It means that it is the same information supposed to always reliably to serve a specific function. For example, in $A \heartsuit B$, the message “I love you” is always supposed to serve the function that a person employs to express her/his love to other people. This makes communication as an interpersonal enterprise be possible. Why does a message stabilizedly and reliably realize a certain function? A user of information may employ (interpret) a physical token as a *Str* being about anything since anything associates with anything in one way or another. However, not any way of employing the physical token, or any *Ref*, serves a certain function that contributes to the interest of the user. Only those *Ref* that contributed to the function in the past is stabilized as the *Nor* of the *Str* (Millikan 1984b; Neander 2017; Shea 2018). As a result, the *Ref* of a *Str* is determined in the stabilizing process. Moreover, emergence of *Nor* as stabilizing processes of the *Ref* of *Strs* has taken the existence of *Ref* and *Str* as premises. It is the stabilized relationships between *Ref* and *Strs*, and the resulted effects that define *Nor*. This is what (b) claims. Moreover, as we can see, *Nor* presupposes *Ref*, which presupposes *Str*. Thus, *Nor* hierarchically depends on *Ref*, which depends on *Str*.

Furthermore, *Nor* concerns with the effect of information. It implies that information has causal power. The causal power of *Nor* as a difference maker does not derive from its physical realizer. Wiener claims that, “Information is information, not matter or energy” (Wiener 1948/1961, 132). The claim is raised at the end of Chapter 5 of his magnum opus, *Computing Machine and Nervous System*. Wiener recognizes that any mechanism that processes information must cost a certain amount of energy, no matter whether it is a computer or a brain.⁹ However, the physical consequences caused by information cannot be explained by the energy cost. Therefore, there is no intrinsic relationship between *Nor* and its physical embodiment.

How to understand the contextual sensitivity of *Nor* regarding the proper function of a message is stabilizing and reliable? I think the crucial reason lies in a confusion between two different levels of *Nor*. *Nor* can be understood as either a type or a token under different conditions: (a) and (b) in the definition of *Nor* are at different levels. When *Nor* is in (a), it refers to its proper function and thus being a type. It is a token in (b) when realizing its proper function (Nor_1) in actual contexts. *Nor* can be seen as a token in two senses: First, what an individual intends to do with a message (Nor_2), and second, the actually realized consequences after receivers receive the message and take actions accordingly (Nor_3). Taking $A \heartsuit B$ as an example, the Nor_2 of the message for *A* is to let *B* learns that he loves him or her. If the message is delivered successfully with no distortion and *B* does not know that *A* loves her before they receive the message, then the Nor_3 is that *B* learns a new knowledge that *A* loves her. In this case, Nor_1 , Nor_2 and Nor_3 are consistent. But there are cases in which these three are inconsistent. For instance, if *A* actually does not love *B* but intends to play with her feelings, then the Nor_2 of the message mismatches the Nor_1 . If the telegrapher at the decoding end wrongly decodes the message into a German one, “Ich

Liebe Dich,” and *B* happens to know no German, then the *Nor*₃ for *B* will be that *A* sends meaningless symbol sequences. That is to say, *Nor*₃ is inconsistent with *Nor*₁ and *Nor*₂.

If *Nor* is understood at these two different levels, there will still be some remaining questions: What are the relationships between *Nor*₁, *Nor*₂ and *Nor*₃? How does *Nor*₁ as a type emerge? Does it emerge from *Nor*₂ and *Nor*₃, namely a type from tokens? In turn, how is *Nor*₁ realized to be *Nor*₂ and *Nor*₃, that is, type to token? An underlying metaphysical question is: Can non-physical entities have physical consequences?

These questions intimately connect with the theorem of double contingency in communication (Vanderstraeten 2002), first proposed by Talcott Parsons (1951) and further investigated by Niklas Luhmann (1984). The theorem proposes a conceptual possibility that miscommunication happens in social interactions. Suppose two human agents, Amy and Billy, interact with each other through a set of signals. For Amy, as a part of the communication, she has to choose a signal that properly represents her idea and successfully achieves her intention. While for Billy, that is, the other aspect of communication, he has the freedom to choose his own way to understand Amy’s signal. The double contingency in the example refers to the freedom of both parties to make alternative selections of her/his own interpretation. That is to say, how Amy and Billy understand the signal is contingent. It is possible that the ways are different: Amy’s understanding is different from Billy’s. As a result, miscommunication can occur. However, communication is normal, while miscommunication is rare in social life. Why?

It is clearer to illustrate the theorem of double contingency through *Nor*. The ways Amy and Billy personally understand the signal are *Nor*₂s for each. Respectively, the result of one’s *Nor*₂ is the *Nor*₃ for the other. In the case of miscommunication, their *Nor*₁s are different from each other and, as a result, *Nor*₁ is inconsistent with *Nor*₃ for each. Communication is normal while miscommunication is rare because there is *Nor*₁ of the signal. For each, Amy’s and Billy’s *Nor*₂ and *Nor*₃s are consistent with *Nor*₁. Thus, the problem of communication is translated to the questions mentioned above. However, providing answers to those questions is the project of theory construction rather than of conceptual analysis.

As far as I know, compared with the discussion on how to measure the quantity of *Str* and *Ref*, there is relatively less discussion on the measurability of *Nor*. Mark Burgin (2010) is an exception. He reviews three kinds of theories of *Nor*: Economic, mission-oriented and transformational approaches. A natural way to think of *Nor* is to consider the value of information has in decision-making and action. *Nor* is determined by its contribution to one’s decision-making and the outcome of her action according to the decision. It can be measured by the changes of the probability distributions of one’s expectations and of the effects of actions with respect to the intended goal (Burgin 2010, 414). It is easy to make correspondences between the theories and the terms I employ in the paper: Changes of expectation correspond to *Nor*₂, the actual effects of actions to *Nor*₃ and the stabilizing probability distribution of those two to *Nor*₁.

Moreover, it is necessary to note that Nor_1 is not the average of Nor_2 s and Nor_3 s as Millikan (1984a; 2004; 2005) has emphasized. The theories reviewed by Burgin are about the average of Nor_2 and Nor_3 . Therefore, a formal theory of Nor_3 is still in need. Signaling games theory first proposed by David Lewis (1969) and further developed by Brian Skyrms (2010a) may fill the vacuum. Signaling game theory aims to explain how the meaning (Nor) of signals spontaneously emerges and evolves in social interactions between senders and receivers.

Another way to formally characterize Nor was proposed by Charles Sanders Peirce which has been rediscovered by researchers today (DE Tienne 2005; Nöth 2013; Liszka 2016). Different from those theories formally characterizing information through calculus of probability, Peirce uses logical quantities to measure Nor . The theory measures information contributing to the growth of knowledge through measuring the changes of the breadth and depth of knowledge. I think that, despite of the methods employed, the theory is not incoherent with those theories discussed above if we understand knowledge as justified true beliefs, which are mental states resulted of being informed.

In the conceptual analysis of the structural, referential and normative aspects of information, it can be recognized that these three different aspects are not just different characterizations from different perspectives. As a result, the analysis also reveals the hierarchically nested nature of these three aspects that they are asymmetrically dependent on one another. Nor depends on Ref , which depends on Str . In turn, an analysis of Str can provisionally ignore the consideration of Ref and Nor ; an analysis of Ref requires consideration of Str but can provisionally ignore the consideration of Nor , while an analysis of Nor requires consideration of both Ref and Str . In summary, Str , Ref and Nor are hierarchically dependent. The analysis of the hierarchy of Str , Ref and Nor is in the same spirit of Bates (2005; 2006) and Deacon (2006; 2007), though the approach employed here is different from theirs. Because of the hierarchical nature of information, the concept of information can be used in different domains at proper levels of abstraction, as Floridi (2011) has argued.

2.6 Conclusion

Information is something carried by *signals about* something *for* some use. Theorists abandoned the aboutness and usefulness of information for engineering purposes in early mathematical theories. Nevertheless, the usage of the concept of information in many fields still unavoidably involves those aspects. It leads to much confusion regarding conflating information with many other concepts, like messages, data, computing, codes and meaning. In order to understand information, two kinds of work are necessary: Conceptual analysis and theory construction. Theory construction should take conceptual analysis as a premise because the latter provides a map upon which problems of information are explicitly marked. The chapter sought to make a conceptual analysis of information.

Although there is no agreement on what information is, it seems that people do believe that information is a tripartite phenomenon. I adopt Terrence Deacon's

terms to refer to those three aspects of information: Structural, referential and normative. Structural information originates from symmetry breaking. Where there is difference, there is structural information. Structural information is objective and thus can be measured intrinsically and extrinsically. Referential information concerns the referential relationship between signal and source. It should not be confused with meaning—which occupies a central place in the philosophy of language. Referential information is much more basic than meaning and not limited to linguistic phenomena. It mismatches with meaning and is alethically neutral. The normative aspect is the usefulness or function of information. It can be understood at two levels: Type and token. A signal normally has stabilized functions. The stabilized function is a type which may be different from intentional function and actually realized function. Structural, referential and normative information are hierarchically nested that the normative depends on the referential, which depends on the structural.

With the conceptual analysis, it is possible to construct a theory of information. I believe that the theory of information we expect should explain structural, referential and normative information, while the most successful theories, aka Shannon's mathematical theory of information and algorithm theory of information, as we have seen in the paper, are concerned only with the measurement of the quantity of structural information. Since then, however, meaning has been lost in many formal theories. Therefore, a task of information studies today is to restore referential and normative aspects in the understanding of the concept of information.

Notes

- 1 By no means everyone will agree with the folk idea that information is a tripartite phenomenon. Many people, especially those who work in science, take Shannon's view of information as their intuition. A motivation, if we understand correctly, behind the movement of pursuing a unified theory of information is to go beyond Shannon's mathematical theory of communication. Other people may believe that information has something to do with people. The common sense used in this chapter is inferred from the imagined daily scenario taken from Yehoshua Bar-Hillel (1955). Instead of uncritically taking it as the basis for the conceptual analysis of information, we use it as a start and then to specify the set of conditions that can define each aspect of information.
- 2 The confusion between information and words (structural aspect of information) will be articulated in [Section 2.3](#), and the confusion between meaning and information will be cleared in [Section 2.4](#).
- 3 Dretske (1981) has argued that causal relations are neither necessary nor sufficient conditions for informational relations. I agree with his argument that detailed description of the causal process upon which an informational relation depends cannot reveal what the information is. However, I do not think his argument that causal relation is not necessary for informational relation is successful since his interpretation of causality is dubious. He defines causal relation as a regular, law-like succession between two events. He gives two cases in the book to argue that it is an informational relation but not causal. One is that an event may indeterminately cause other three different events in equal probability. The other is that an event causes two events. Although there is no causal relation between the caused events, any of the two can be informed by each other. These two cases are not causal relations according to the definition. I think his definition is too strong to include periphery cases of causation. According to some recent works on

causation, there is no problems to include the two cases as causal relations (See the entry “probabilistic causation,” in Stanford Encyclopedia of Philosophy, for instance). In order to avoid unnecessary debates, I use a weaker concept, “physical relation”, rather than “causal relation.”

- 4 The difficulty is called content-determinacy challenge is raised by [Fodor \(1990\)](#) in arguing against the teleosemantic programme. It is also mentioned by Dretske himself (1981). A detailed formulation of the difficulty can see Chapter 7 of [Neander \(2017\)](#).
- 5 The argument below comes from ([Dretske 1981](#), 41–47). However, I disagree with him that “[t]he information embodied in a signal (linguistic or otherwise) is only incidentally related to the meaning (if any) of that signal.” I believe the relation between the information and meaning of a sign is not just incidental.
- 6 Generally, the discussions of disinformation and misinformation concern with action. See [Fox \(1983\)](#) and [Fallis \(2015\)](#).
- 7 [Floridi \(2011, Chapter 4\)](#) gives two arguments on the Veridicality Thesis. The first argument is that the use of “false” in “false information” is attributive. There will be semantic loss if we split “false” from the whole phrase. While the use of “true” in “true information” is predicative that without which there is no semantic loss. The second one is a semantic argument verifying that it will be hard to understand ordinary phenomena of semantic erosion.
- 8 Some scholars argue that the so-called Bar-Hillel/Carnap Paradox is superficial, see [Fresco and Michael \(2016\)](#), and Floridi confuses information with informativeness, see [Duzi \(2010\)](#). Even their arguments are correct, I believe that Floridi’s theory of strongly semantic information has its own value for its supposed function.
- 9 This is called *Landaur principle* that a bit of information requires at least $kT \ln 2$ energy. See R. [Landauer \(1961\)](#).

3 Information is physical (negatively)

3.1 Introduction

To naturalize intentionality or aboutness, many philosophers believe it must be grounded in nature.¹ That is, the relationship of aboutness must be based on some kind of natural correlation. Since Paul [Grice \(1957\)](#), natural information has often been cited as a candidate. Thus, the task for philosophers becomes one of understanding what natural information is. [Dretske \(1981; 1988\)](#) views it as nomic regularities, while [Millikan \(1984; 2004; 2017\)](#) and others (e.g., [Stegmann 2015; Scarantino 2015; Shea 2018](#)) locate it as correlational information.²

Taking Millikan's conception of infosigns as an exemplar—since her theory is among the most sophisticated—“an infosign is, first and foremost, a member of an infosign-infosigned pair that exemplifies a non-accidental correlation between signs and states of affairs, with the signs all corresponding to the infosignified states of affairs according to the same projection rules” within a local domain ([Millikan 2017](#), 110). Infosign and infosigned are types of structured world affairs. We may treat each part of the pair as two sets of world affairs. The non-accidental correlation between the two sets represents an isomorphic mapping in a mathematical sense. That is, a property of one affair projects onto a property of the other according to the same mapping rules. Together, we may say that one set is structurally isomorphic to the other. In other words, the form or pattern of one type maps isomorphically to that of the other. Therefore, we can say that an infosign re-presents the form of its infosigned according to the same projection rules.

Let's illustrate this concept of correlational information, or the infosign-infosignified pair, using [Millikan's \(2004, 47–51\)](#) example. Normally, a dark cloud serves as an infosign of rain. Statistically, given the presence of a dark cloud, the probability of rain increases. However, according to Millikan, “strictly speaking, it is not the black cloud that is a sign of rain” ([Millikan 2004](#), 47). A black cloud and rain belong to two distinct sets of world affairs. Members in each set share the same structure or form. The structure of the black cloud set serves as an infosign for the structure of the rain set. There are isomorphic mapping rules between the two sets, according to which one structure projects onto the other. Thus, we may say that the form of rain is projected onto the form of the cloud according to the mapping rules. The form of the cloud re-presents the form of

the rain. When an observer perceives the cloud, the form of rain is communicated to them.

In this way, Millikan's conception of correlational information, or the infosign-infosignified pair, resembles the communication of form within Shannon's framework of communication systems. What originates at the source is form. This form is carried by an infosign and communicated to the receiver. The receiver decodes the form of the event at the source, carried by the infosign, according to the projection rules. However, this formulation raises a metaphysical question. The forms are shared by members of both sets and are, thus, abstract. So, where does abstractness fit within nature? What is the relationship between abstract form and particular physical events? This question is closely tied to another issue in the philosophy of information: What exactly is information? What is the metaphysical status of information?

The contemporary metaphysical problem of information can be traced back to Norbert Wiener's (1961, 132) famous slogan: "Information is information, not matter or energy." This claim appears in the final chapter of *The Human Use of Human Beings: Cybernetics and Society*. Wiener acknowledges that any physical mechanism processing information must expend energy, whether it's a computer or the brain. However, he asserts that the energy cost is not an accurate measure of the effect produced by a piece of information. Therefore, information cannot be identified with the physical processes that implement it. Wiener's insight deepens the puzzle: How can something non-physical have physical effects? Where does the causal power of form over the physical come from?

The debates over the metaphysical status of form date back to Plato and Aristotle. In contemporary information studies, there is still disagreement about its interpretation. Some argue that forms or patterns are nothing more than the organization or structure of world affairs—regularities in nature. These are not something other than the physical. We may call this view "reducible physicalism." This position is adopted by most philosophers in the naturalistic tradition of meaning, such as Stampe's causal theory (1977), Dretske's informational semantics (1981; 1988), Millikan's biosemantics (1984; 2004; 2017) and informational teleosemantics (e.g., Neander 2017; Shea 2018). As I argued above, however, these theories still leave us with an explanation of the place of form in nature.

Another position, called "fundamentalism," treats information as a primitive and fundamental property, as basic as matter or energy (Chalmers 1996). According to this view, information is identical in form, pattern, difference, structure, organization, etc. As long as there is a difference or symmetry breaking, there is information. In this sense, the universe is fundamentally computable and thus digital (Zuse 1967; 1969; Schmidhuber 1997; Wolfram 2002; Dodig-Crnkovic 2011, among others). Furthermore, information cannot be reduced to the physical; instead, as John Archibald Wheeler (1989) famously put it, "It from bit." At the most fundamental level, the universe is not made of matter or energy, but of information. Epistemologically, this view holds that information is not an explanandum but an explanans. Since information is a basic

property of the universe, we need not explain it; rather, we can use it to explain other phenomena, such as cognition, perception, consciousness and even matter itself.

Most people, however, reject this position. Methodologically, it is often criticized as an “ostrich policy,” in that it avoids addressing the core problem by claiming that information is primitive without offering a proper explanation. Epistemologically, it conflates two levels of explanation. The fact that an event is computable does not imply it is fundamentally digital. As [Floridi \(2011\)](#) argues, these theories are lacking in specific boundary conditions under which they are applicable, making them empty. Moreover, treating information as synonymous with terms like “organization,” “difference” or “structure” fails to explain what information is in common sense. Ontologically, this view is a kind of Platonic realism, which can lead to the same metaphysical difficulties faced by Platonism since Aristotle.

Although reducible physicalism and fundamentalism are in conflict, they share a common way of thinking: They treat the concept of form positively. That is, they think of form as something extra to the physical and attempt to discuss the possible relationship between them. This dualistic approach is useful for considering the possible relationship between different concepts, but it can also be risky, leading to philosophical pitfalls. For instance, if a form is not physical and is even abstract, as [Millikan \(2004\)](#) and others argue, how can it coexist with the physical, which exists in particular spatial-temporal locations? How do the physical and non-physical interact? These are the classical difficulties faced by dualism.

[Deacon \(2007; 2008; 2012a; 2017\)](#) proposes a negative view of form. Briefly, he argues that form is an intrinsically absent aspect of something present. This negative understanding of form not only solves the difficulties in explaining the concept of form but also bridges informational and thermodynamic entropies, laying a solid foundation for understanding the physical basis of information, representation, reference, meaning, or intentionality. In this chapter, I will reformulate this view and use it to explain the physicality of information, and by extension, intentionality.

In [Section 3.2](#), I argue that understanding information as the communication of form revives its etymological meaning. Therefore, the key to understanding information lies in understanding form. Deacon proposes that form is an intrinsically absent aspect of something present. This view helps us resolve a longstanding puzzle in information studies: The superficial similarity between the formulas for entropy in Shannon’s theory of communication and thermodynamics. It reveals the physicality of information and establishes a foundation for understanding reference and intentionality, as will be discussed in [Sections 3.3](#) and [3.4](#). However, in [Section 3.5](#), unlike fundamentalism, I argue that the physical ground of information is not equivalent to information itself, or we risk encountering problems such as misinformation and content indeterminacy. [Section 3.6](#) concludes the chapter.

3.2 The place of form in nature

The most successful and widely used theory of information, Shannon’s mathematical theory of communication, ultimately loses the very meaning of information.

To revive the original meaning of this concept, the first step is to characterize it. A productive way to do this is by examining its etymology, as Austin suggests,

[A] word never-well, hardly ever-shakes off its etymology and its formation. In spite of all changes in the extensions of and additions to its meanings, and indeed rather pervading and governing these, there will still persist the old idea ... Going back into the history of a word, very often into Latin, we come back pretty commonly to pictures or models of how things happen or are done.

(Austin 1961, 149–150. In [Capurro & Hjøland, 2003](#))

Intuitively, the term “information” consists of three parts: *in*, *form* and *-ation*. The prefix “in” has the meaning that accentuating an action, as analyzed by [Capurro \(2009\)](#). The suffix “-ation” refers to “the action or process of doing something.” Therefore, the superficial meaning of information is “the action of giving form to something.” This meaning is closely connected to its original Greek-Latin origin—*informatio*. There are two basic meanings of *informatio*: (1) “the action of giving a form to something material” (referred to as AM) and (2) “the act of communicating knowledge to another person” (referred to as AP). These two meanings are closely related. In other words, “the act of communicating knowledge” is the act of transmitting form to a mind to shape it into a certain structure. AM represents the ontological aspect of information, while AP represents the epistemological. There is also a pedagogical aspect derived from AP, which Tertullian (ca. 160–220 A.D.) referred to as *Moses populi informatory*, meaning an educator or guide who shapes or molds young people into true human beings ([Capurro & Hjøland, 2003](#); [Capurro 2009](#)). Theological uses of this Latin origin evolved from these two meanings: One relates to God shaping nature by giving forms to it (i.e., creation), and the other to God educating mortals by providing them with true knowledge ([Hofkirchner 2013](#), 6).

As we can see, those uses of the Greek-Latin origins of the term “information” in epistemology, pedagogy and theology all have their roots in its ontological meaning. [Capurro and Hjøland \(2003\)](#) argue that these two meanings derive from two everyday Greek experiences: Pottery and perception. In pottery, the artisan shapes clay into a specific form with a concept of the desired form in mind. For example, an artisan molds clay into a bowl according to the form of a bowl they have in mind. In the case of perception, we sensually perceive an object—not its substantial aspect but its form. The form is given to shape our minds through perception, and we acquire knowledge through this process of form-giving. This idea was inherited by later thinkers in the Middle and Modern Ages, as we will see later.

“The action of giving form to something” has two implications: The act of giving form and the accomplishment of the act. Using pottery as an example, the procedure of pottery involves at least three steps: (1) a blueprint of the desired form, whether physically printed or mentally conceptualized; (2) the act of shaping the clay and (3) the completion of the final object. In the process, the artisan brings the form of a bowl into the clay to create the bowl. The blueprint must be the correct form of the bowl; if it is not, the artisan will not create the intended object. The

artisan must transform the form into clay accurately, shaping it into the desired bowl. Only when the bowl is finished can we say the artisan has successfully transferred the form of the bowl into the clay, producing the actual object. This etymological understanding of information aligns with Peirce's idea that information is the communication of form. In semiotic processes, what is transmitted from an object through a sign to a receiver's interpretation is form.

This is also a common understanding today. In his classic work "Form, Substance, and Difference," Gregory Bateson articulates this idea by stating, "what we mean by information—the elementary unit of information—is a difference that makes a difference" (Bateson 1972, 478). Before Bateson, Donald MacKay (1953) made a similar statement: "Information is a distinction that makes a difference." Floridi's definition of data shares a similar flavor, as we argued in Chapter 2: "Dd datum =def. x being distinct from y, where x and y are two uninterpreted variables, and the domain is left open to further interpretation" (Floridi 2011, 85). Therefore, form lies at the very core of information. But what exactly is form or difference?

Bateson also provides an answer to this question. He discusses a statement made by Alfred Korzybski, the father of general semantics: "A map is not the territory it represents, but if correct, it has a similar structure to the territory, which accounts for its usefulness." Bateson then asks, "What is it in the territory that gets onto the map?" His answer is that it is certainly not the territory itself that gets onto the map, but rather the differences: "What gets onto the map, in fact, are differences—whether differences in altitude, vegetation, population structure, surface characteristics, or other factors. Differences are the things that get onto the map" (Bateson 1972, 480).

What, then, is a difference? Bateson suggests that a difference is an abstract concept. However, this leads to a conceptual difficulty, as Cashman (2008, 47) argues: If a difference is abstract, it does not exist in space and time. But whatever exists in the territory must exist in space and time. Therefore, there seems to be no actual difference in the territory that can get onto the map. Bateson, thus, faces the same metaphysical dilemma that Millikan (2004) encounters.

Bateson's solution is interpretation. He argues that form and difference derive from the knower. It is the knower's mind that interprets changes in the states of world affairs as form or difference. Therefore, form is an observer-dependent concept. This solution brings us back to the starting point of our investigation. We aimed to ground intentionality or reference in nature, but the foundation we find—form and difference—exists in the observer's mind, not in nature itself. We thus encounter a dilemma: Is form an abstract entity in the Platonic sense, or is it observer-dependent? The former leads to ontological extravagance, which is difficult to accept, while the latter results in a vicious circle. As Deacon (2012a, 189) argues,

To even perceive regularity or pattern this act of observation must itself be grounded in a habit of mind, so to speak. In other words, to attribute physical regularity to some perceived or measured phenomenon presumes a prior mental regularity or habit with respect to which the physical regularity is

assessed. So, unless we grant mentality some other undefined and mysterious means of observational abstraction, claiming that regularities, similarities, and general types are only in the mind but not in the world simply pass the buck, so to speak. Some general tendency must already exist in order to attribute a general tendency to something else.

To avoid the pitfalls in understanding form and difference, one must first ask: Is form an epistemic concept, existing only in the mind, or does it also exist in nature?

Deacon (2012a) suggests that form does indeed exist in nature, but in a negative way. To illustrate this, let's consider the question: How do we define the individuality of a whirlpool in a stream? First, it cannot be defined by the water molecules that make up the whirlpool. These molecules do not stay within the whirlpool for long—they are constantly being replaced. Second, the individual motion of the water molecules does not form a regular pattern; if we observe a single molecule, its path appears irregular. Third, there is no constant, unchanging pattern controlling the motion of the water molecules in the whirlpool. Although we can easily distinguish a whirlpool from its surroundings because of its rough circular symmetry and the general rotation of water, these patterns themselves are not static. The circular symmetry and the center of rotation change over time, and there is no single, unchanging pattern that governs the water's flow. So, what defines the individuality of a whirlpool?

It is the similarity—a similarity that ignores nearly all differences at some level of detail—that allows us to consider the rotating flow of water at different moments as the same whirlpool, suggested by Deacon. This concept of similarity is generalizable. For example, we can define both whirlpools in water and galaxies (which are vastly different in scale) as belonging to the same type, because both exhibit a spiraling shape, even though they differ in constituents, size and location. The spiraling shape they share is abstract and disregards their concrete differences.

This case implies two things: First, form is not tied to the physical properties of the individual elements that make up the system, so it cannot be reduced to these elements. Second, there is no eternal form in the Platonic sense, as this would lead to paradoxes of infiniteness, as criticized by Aristotle. The form, then, is not an intrinsic essence or eternal pattern, but rather must be the result of mental abstraction—a descriptive feature imposed through external analysis. However, this explanation falls into the trap of relying on external measurement, which we sought to avoid from the outset.

Deacon proposes the concept of constraint as a way to sidestep this dilemma due to its negative logic. "In statistical mechanics, constraint is a technical term used to describe a reduction in the degrees of freedom of a system, a restriction on the possible variation of properties" (Deacon 2012a, 192). In simple terms, when a system is constrained, the number of states it can realize is reduced. Constraint does not focus on what is added or present, but on what is excluded or reduced. For example, melted wax could take any form until it is imprinted by a signet. Once the wax is imprinted, it is shaped into a specific structure, and all other possible shapes are eliminated. In this way, the signet reduces the state space of the

wax. The realized form is not determined by some universal, abstract form, but by the concrete possibilities that are excluded by the constraint. The logic behind constraint is that, if not all possible states are realized, the variety in how things can differ is reduced.

This negative logic avoids presupposing extrinsic observation by discussing the realization of form transmission in terms of quantity rather than quality. “The term ‘constraint’ thus denotes the property of being restricted or less variable than possible, all other things being equal, and irrespective of why it is restricted” (Deacon 2012a, 193), since there are multiple ways to realize the same constraint. Therefore, “employing the concept of constraint instead of the concept of organization (such as pattern, order, form, difference, etc., *added by the author*) not only avoids observer-dependent criteria for distinguishing patterns, but also undermines value-laden notions of what is orderly and what is not” (Deacon 2012a, 195).

In the example of the signet and the wax, the imprinting process can be described as the transmission of form from the signet to the wax. This implies that it is the form determining the shape or arrangement of the wax. The form is qualitatively different from the concrete physical state of the wax because the form is abstract and universal. Form can only be intentionally discriminated by the agent who has interpretive competence because it is abstract that does not exist in concrete particulars. However, if we think in terms of negative logic, the change in the wax is not qualitative, but quantitative. Assume there are n states that the melted wax can potentially assume before being imprinted by the signet. After being imprinted, $n - 1$ states are reduced, leaving only one state. In this way, the reduced states have extension, as they are contrasted with what is present and thus are concrete, existing in space and time, albeit in a negative way. As a result, a constraint also exists in space and time, defined by the manifested states and those that would have been realized but were eliminated. Therefore, the concept of constraint avoids the difficulties encountered by Bateson’s conception of form. Additionally, the presupposition of extrinsic observation is avoided, and the form can be described in dynamic terms through transformed into the concept of constraint.

Peirce’s idea of information as the communication of form, and the contemporary notion of information as a distinction that makes a difference, can also be reformulated using the concept of constraint. The communication of form is constraint propagation. Consider Bateson’s statement, “Where there is no difference, it can cause no difference.” It means that “what something doesn’t exhibit, it can’t impose on something else via interaction” (Deacon 2012a, 198). Just as constraint reduces the possibilities of a system, it indirectly determines which differences can and cannot make a difference in interactions.

This has two complementary consequences. Whenever existing variations are suppressed or otherwise prevented from making a difference in any interaction, they cannot be a source of causal influence; but whenever new constraints are generated, a specific capacity to do work is also generated.

(Deacon 2012a, 198–199)

In line with this framework, the semantic mapping rules within Millikan's bio-semantics can be understood not as abstract representations, but as constraints. In this view, what transfers from an event to an infosign are the constraints imposed by the event. That is, the constraints inherent in the event propagate through the sign, which then embodies these constraints. This perspective helps avoid some of the metaphysical difficulties that Millikan's theory encounters by grounding meaning in natural constraints rather than abstract mappings or external forms.

Moreover, this approach aligns with the etymological meaning of information, where the "action of giving a form to something material" can be interpreted as the process of propagating constraints from the source to the receiver through the sign. The constraints originating from the source shape the state of the receiver. This negative conception of form—as constraints—also provides an insightful framework to reconsider the physicality of form communication, especially in relation to the concepts of thermodynamic and informational entropies. By focusing on constraints, we ground the communication of information in the physical dynamics of systems, thereby revealing the physical basis for both information and reference.

3.3 Bridging the formal and the physical through entropies

Naturalistic philosophers often treat correlational information, the natural grounding of intentionality, as statistical. They follow Shannon's steps, using probability to characterize it.³ The probabilistic understanding of information actually implies a negative logic. For example, according to Shannon's mathematical theory of communication, information is the reduction in uncertainty. As I briefly discussed in [Chapter 2](#), the quantity of information carried by a signal is not determined by the intrinsic properties of the signal itself, but by the set of possible signals that could have been sent. In other words, the information carried by a signal is determined not by what is present, but by the signals that are absent.

Without reference to this absent background of possible alternatives, the amount of potential information of a message cannot be measured. In other words, the background of un-chosen signals is a critical determinant of what makes the received signals capable of conveying information. No alternatives = no uncertainty = no information.

([Deacon 2012a](#), 134)

Let's further illustrate this idea with an example.

Suppose there is a world affair at the source, which has four possible states—(S_1 , S_2 , S_3 , S_4)—with equal probability, and four possible signals—(M_1 , M_2 , M_3 , M_4). These signals are a set of world affairs that recurrently correlate with the world affair at the source, with each signal corresponding to one state: M_1 appears when S_1 occurs, M_2 when S_2 , M_3 when S_3 and M_4 when S_4 . When S_1 occurs, M_1 following appears and carries the information of S_1 . Assuming a noiseless channel, the quantity of information produced at the source and carried by M_1 is the same—2 bits. This information is determined by those states that could have occurred but

are actually not present. For the receiver, the quantity of information is determined by the three possible signals that could have been sent but are absent. We can understand the informational content in the same way.

As I reviewed in [Chapter 1](#), one sophisticated way to characterize informational content is to formulate it as vectors—the probabilistic distribution of all possible states of the world affair at the source. The probabilistic distribution of all the possible states of the system can be seen as a formal characterization of the constraints the system imposes. When a system has some constraint, it means that some possible states cannot be realized, and as a result, some states do appear. Taking the probabilistic difference-maker theory as an example: According to the theory, a full description of the informational content of a signal includes two parts—incremental information and the overall support conferred by the signal (shortened to overall information) ([Scarantino 2015](#)). Incremental information characterizes the changes in the probabilistic distribution of the world affairs, while overall information describes the posterior probabilistic distribution. In terms of constraints, incremental information reflects changes in the constraint, while overall information describes the resulting constraint. With respect to the mapping rules between the sets of possible states and signals, the constraints and their changes also project onto the signals. That is, the constraints of the world affairs are propagated through the signals.

From this understanding, we can infer two points: First, natural information, as a formal property of world affairs, can be seen as a characterization of constraint. It exists in space and time, realized by the physical states of the world affairs at the source and the signals. Second, when there is incremental information, something diverges from the original state of the world affairs at the source, resulting in changes to the original constraint. When something changes, in turn, it produces information, “information is a difference which makes a difference.” This negative understanding lays a foundation for understanding how correlational information in nature, as formal properties, is realized physically.

A good starting point to understand the physicality of information is the dual nature of signals. On the one hand, we measure the quantity of information based on the formal properties of signals; on the other hand, since the form of world affairs is the constraint realized by the physical states, signals follow the Second Law of Thermodynamics. Therefore, a signal can be characterized by two types of entropy: Informational entropy and thermodynamic entropy. “The Shannon entropy of a signal is the probability of receiving a given signal among those possible, and the Boltzmann entropy of the signal is the probability that a given signal may have been corrupted” ([Deacon 2007](#), 135). Furthermore, the two entropies share the same formula, suggesting a relationship between them. Since the inception of Shannon’s theory of communication, scholars have been searching for this relationship.

At the beginning of information studies, a story circulated that Shannon could not find a proper name for the equation he proposed to measure the quantity of information. When he visited Von Neumann and expressed his confusion, Von Neumann said, “You should call it entropy, for two reasons. First,

your uncertainty function has been used in statistical mechanics under that name, so it already has a name. Second, and more importantly, nobody knows what entropy really is, so in a debate, you will always have the advantage.”⁴ One thing should be clear: Despite sharing the same formula, the two entropies belong to different logical categories and cannot simply be equated. Information is formal and logical, while thermodynamics is physical. Moreover, thermodynamic entropy of a system tends to increase when the system is isolated, while information entropy does not generally increase spontaneously in a communication system (Deacon 2007, 135). So, what is the relationship between thermodynamic entropy and Shannon’s information theory? Are they merely accidentally similar, or is there a deeper connection between them beneath their superficial similarity?⁵ To investigate this, we first need to clarify what thermodynamic entropy is.

Thermodynamic entropy measures the disorder of a system. It can also be understood from the perspective of constraint, as discussed in the previous section. According to the Second Law of Thermodynamics, a thermodynamic system tends toward equilibrium. In terms of entropy, it tends to become more disordered, with its elements becoming more and more uncorrelated until they reach an even distribution—maximum entropy. As a result, each element becomes less constrained, gaining more freedom. This means that the system could be in any possible state, but the probability of being in any specific state is astronomically low because the elements are evenly distributed and uncorrelated. When this tendency is prevented or reversed in a system—that is, when elements tend to correlate locally—it suggests that something external is disturbing the system. Negatively speaking, when a highly probable state fails to occur, it typically indicates external influence. In other words, the probabilistic distribution of all possible states changes due to external influences. This connects thermodynamic entropy to information.

The interactions between the system and the external world can be seen as a communication system. The possible states of the system can be used as signals, providing information about the world affairs acting upon it. The change in the probabilistic distribution can be measured as incremental information. In turn, information depends on the physical features of the system serving as the channel and the signals. On the one hand, all possible states of the system can be employed as informational states; on the other hand, the number of possible states determines the maximum information capacity. Furthermore, as I argued earlier, information entropy does not have a natural tendency to increase like thermodynamic entropy. However, when considering the physical medium that information embodies, things change. Since informational states are realized by and mapped onto the physical states of a thermodynamic system, the tendency to increase entropy spontaneously corrupts the channel and signals, reducing the system’s information capacity and making the channel less reliable.

Now that we have articulated the relationship between thermodynamics and informational entropies, let’s examine how this relationship helps clarify the natural grounding of representation.

3.4 Grounding re-presentation in the physical

According to many naturalistic philosophers, the correlations between events in nature serve as the natural ground for representation. When a world affair recurrently correlates with another (in a statistical sense) within a local domain, the natural information of the former is transmitted to the latter. In terms of constraint, the constraints of the former propagate constraints in the latter. This provides a formal characterization of natural information. But how do we understand its physicality? We can approach this question by following Shannon's work and building on Deacon's ideas (2007; 2017).

A key lesson from Shannon is that the information conveyed by a signal is not determined by the intrinsic properties of the signal itself. Rather, it is the physical properties that make representation or reference possible. But how can the physical properties of signals possess this capacity? A clue was given in the previous section: Signals, as physical media, can be analyzed in two dimensions—formal and physical—with the formal aspect depending on the physical one. This implies that information transmission must be a physical process and, therefore, is subject to the laws of thermodynamics.

So, the basis for the interdependence of Shannon and Boltzmann entropy can be stated in simple form as follows: *a reduction of either Shannon or Boltzmann entropy does not tend to occur spontaneously, so when it does occur it is evidence of the intervention of an external influence.*

(Deacon 2007, 138)

Let us consider a thermodynamic system, T . When there is no external influence, the entropy of T spontaneously increases until it reaches its maximum. When entropy stops increasing or even reverses, with elements tending to correlate with each other—thus forming structure or order out of chaos—something external to the system must be disturbing it and changing the tendency. Therefore, when T is in a spontaneously improbable state, it is evidence of an extrinsic disturbance. This is a natural affordance for information entropy. According to Shannon's theory, information is the reduction of uncertainty. The more uncertain a signal is, the more information entropy it carries. The uncertainty of a signal is measured by the number of possible signals. The possible signals correspond to the possible states of a world affair. Thus, a system with more thermodynamic entropy also embodies more information entropy. In other words, it is more uncertain whether the system can be in a specific state. Therefore, information entropy is constrained by thermodynamic entropy. When there is a reduction in the uncertainty of a signal, or information entropy, it means some external constraint has been applied to the signal medium. The signal then carries correlational information about the world affair that imposed the external constraint. This logic applies equally to the opposite case in thermodynamics.

Now, suppose a system is in a highly constrained state far from equilibrium, like a metal detector system (Deacon 2017). When a metal detector operates, a

constant flow of electric current through a coil forms a stable electromagnetic field. The stability of the field means that its highly constrained state persists spontaneously if the detector is turned on and there is no perturbation. The entropy of the system is very low. However, if the detector touches something conductive, the field is easily distorted as the electric current is disturbed. This means the field becomes less constrained, gaining more degrees of freedom and the system's entropy increases. In this case, the spontaneous state of the detector system has low entropy. When this low-entropy state is lost, it indicates the influence of something external. Analogously, this dynamic interaction can also ground information entropy. The electromagnetic field is used as a signal medium. When the system is turned on, the information entropy is low because the possible states that can serve as signals and channels are few. When a conducting object disturbs the field, it adds new possible states, thus creating new channels. This indicates external interference that performs work on the system. Furthermore, we see that the information conveyed by a signal is not determined by the intrinsic properties of the signal itself.

In either case, if we should encounter such a signal it would point us toward the likelihood that this can be informative about something other than the signal detection circuit, something not present in the system itself but rather something outside altering and thereby imposing non-spontaneous constraint on the otherwise uncorrelated jumble of signals.

(Deacon 2007, 138)

This analysis shows that not only thermodynamic entropy grounds information entropy, but also the physical dynamics—according to the physical context—ground representation, or referential information. As I've argued earlier, we may treat a system (whether near equilibrium or far from equilibrium) as a signal medium, with the possible states of the system serving as possible signals, and external events performing work on the system as referent objects. The physical dynamic interactions between the system and the external event can then be employed as the communication channel. The external event's constraints are transmitted to the system, changing the system and forming a new constraint. "In purely physical terms this can be described as a coupling between two systems' states or dynamics so that the behavior of one will partially re-embody some aspect of the regularity or constraints exhibited by the other with respect to their possible modes of interaction" (Deacon 2007, 139). This physical connection forms a natural foundation for correlational information.

Although it is not conventional to cite an entire section of a paper to present one's argument, I believe it is warranted to fully cite Deacon's summary of his argument here. With its step-by-step articulation, it is easy to see the physical grounding of information and, consequently, of representation.

- 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 versus 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. (See notes below.).
- 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 2nd 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 media are physical structures or otherwise passive systems at or near a 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: non-passive 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 by 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.

(Deacon 2017, 10–12)

3.5 The misinformation problem and the content indeterminacy problem

With Deacon’s dynamic theory of the physicality of information, we have addressed the metaphysical status of information and located the natural ground for referential information or representation. Since constraints and constraint propagation are ubiquitous, some argue that this constitutes information, which pervades nature and may even be fundamental to the universe. This is the idea of fundamentalism introduced in the introduction. In this section, I will argue that fundamentalism about information and meaning is mistaken, due to the normativity of information and thus the misinformation problem and the content indeterminacy problem.

Again, Paul [Grice \(1957\)](#) distinguishes between two types of meaning: Natural and non-natural meaning. A key difference is that non-natural meaning can be false, whereas natural meaning cannot. For example, a sign that has the power to be false means something that is not the case. A boy might cry, “The wolf is here!” when there is no wolf, whereas a dark cloud does not indicate rain unless it actually rains. This demonstrates that representation has the power to misrepresent, while correlations in nature *per se* cannot. Nature, simply as it is, cannot be wrong. However, some might argue that when one sees a dark cloud and interprets it as a natural sign of rain, yet no rain follows, the cloud as a natural sign can be false. In this case, the cloud is interpreted as a sign of rain, but it does not itself correlate with rain. The power to misrepresent is essential to understanding intentionality, because intentionality entails the capacity to get things wrong. As Dretske (1986, 65) puts it, “Whatever word we use to describe the relation of interest (representation? meaning?), it is the power to misrepresent, the capacity to get things wrong, to say things that are not true, that helps define the relation of interest.” This is the normative aspect of representation or aboutness. Only with normativity can a representation be true or false, accurate or inaccurate. Normativity is a teleological concept—something can only be said to be true or false with respect to a purpose. Correlations in nature are simply what they are; they are not teleological and, thus, not normative. Therefore, we cannot equate the natural ground of representation, or

locally recurrent correlation, with representation itself. Furthermore, to understand intentionality or meaning, we must explore the source of the normative aspect of representation.

A second challenge fundamentalism faces is the content indeterminacy problem (Dretske 1981; 1986; 1988; Fodor 1984; Neander 1995). A representation typically has definite content—the things represented—but this content cannot be determined merely by the correlation itself. Specifically, for any given sign and its corresponding represented entity, there seems to be more than one way to describe the represented (Neander 1995, 113). This is known as the disjunction problem. To illustrate, let's take the example proposed by Dretske (1981; 1986).

Suppose you hear the sound of a doorbell, which informs you that a visitor is outside the door. In other words, the sound of the doorbell indicates the presence of the visitor. Indicating the visitor is the definite function of the doorbell. The doorbell system is a communication system realized through a sequence of physical events: Pressing the doorbell button by a visitor activates an electrical circuit, which powers an electromagnet, generating a magnetic field, which powers a mechanism that creates a sound, which causes vibration in the eardrum of someone inside the house, ultimately producing neuro-signals sent to the brain and so on. If we think of this process in terms of constraint propagation, we could say that the constraint produced by the pressing of the doorbell button is propagated through the neuro-signals. Through this communication of constraints, the information about the visitor outside the door is transmitted to the person who hears the sound. However, when a person hears the sound, it could carry information about any event in this sequence, as each event in the sequence reliably correlates with the others, and constraints in one event are propagated to others. Thus, we cannot determine the content of the sound simply by examining the physical processes that realize the doorbell.

Additionally, considering the proximity of the events: The visitor pressing the button is a distal event compared to the vibration of the eardrum of the receiver. So why is the sound signal about the distal event, the visitor outside the door, rather than the proximal event? This question cannot be answered purely by constraint propagation. This is a case in which different events causally relate within the same sequence of events (Neander 2017). Given a signal M received by a receiver through a sequence of events, $E_1, E_2, E_3, \dots, E_n$, M could indicate any event in this sequence, as any correlation between M and an event is a process of constraint propagation. So, why is M about the distal event rather than a proximal one? Constraint propagation alone cannot determine the reference of M —it simply happens.

There are also cases where different events co-occur within the same local context. Dretske (1986; 1988) imagines marine bacteria that live in the northern hemisphere and have internal magnets. These magnets align the bacteria to the earth's magnetic field, which inclines downward in the northern hemisphere. As a result, the bacteria propel themselves toward geomagnetic north, moving away from the surface water and into the oxygen-free depths of the ocean, as they can only survive in the absence of oxygen. In this case, the movement toward the ocean depths and the oxygen-free environment co-occur at the same time and place. But what do these magnetic lines refer to—the geomagnetic north or the

oxygen-free environment? Formally, given a signal N , the probabilities of a range of events, E_1, E_2, E_3 , are the same. In other words, N reliably correlates with each of the events in this range. So, which event does N represent? Or why does N represent one event over the others? Locally recurrent correlations simply happen and do not select for one over the other. They cannot fix what a sign refers to. To resolve this, there must be a relation of interest that determines the content of a sign. Therefore, correlations in nature, or constraint propagation, are not sufficient to define meaning. Fundamentalism is mistaken. Following Millikan (2017), we can view constraint propagation as an affordance for referential relationships, but not something that determines meaning on its own. To fully understand meaning, we must investigate where the normative aspect of it originates. This is the topic of the next chapter.

3.6 Conclusion

The physicality of meaning has often been overlooked in discussions about the naturalization of meaning. By appealing to the concept of correlational information, thinkers in the informational and teleological traditions leave an important metaphysical question unaddressed: What is the form in nature? Fundamentalism about information even claims that form, as something abstract, is fundamental to the universe. It is fundamentally different from concrete matter and even grounds the latter. This claim revisits the old metaphysical challenges rooted in the Platonic conception of *eidos*. By reformulating form as constraint, Deacon argues that the form of a world affair is a concrete abstraction, defined by those states that could have been present but are, in fact, absent. This negative logic can also be applied to bridge thermodynamic and informational entropies. With this understanding, we can grasp the physicality of representation as constraint propagation. Peirce's definition of information as the communication of form, and Bateson's conception that information is a difference which makes a difference, can both be reformulated physically through this framework. Representation is grounded in the constraint propagation produced by physical dynamic interactions. However, in contrast to the fundamentalist view of information, constraint propagation grounds meaning but does not determine it. The misrepresentation problem and the content indeterminacy problem show that meaning is inherently normative. Therefore, to provide a comprehensive understanding of meaning, we must understand the normative aspect of information.

Notes

- 1 In this chapter, I do not distinguish concepts of intentionality, aboutness, meaning and representation. So, I use these concepts interchangeably in this chapter unless I give special explanation.
- 2 This family of theories has been reviewed in [Section 1.3, Chapter 1](#).
- 3 People who have interest in knowing more may read [Chapter 1](#).
- 4 http://en.wikipedia.org/wiki/History_of_entropy

- 5 Most researchers who consider the relationship between the two entropies investigate how much energy information costs. They aim to answer a question proposed by James Maxwell. In a famous thought experiment, Maxwell imagines a demon who can use information to violate the almost unavoidable tendency to thermodynamic equilibrium. Latter researchers try to prove that it cannot violate the Second Law of Thermodynamics since the amount of energy cost by a bit of information is more than it reduces through Szilard engine, a thought experiment revised from Maxwell's demon by Leo [Szilard \(1929\)](#). Although this is not unrelated to the topic of this chapter, follow Deacon's footsteps, I will approach the relation between the two entropies in a different way. Those who have interest in the Szilard engine tradition may read the entry "Information processing and thermodynamic entropy" by Owen [Maroney \(2009\)](#).

4 Interpretation

4.1 Introduction

Correlations exist everywhere in nature. Natural correlations can be seen as projective rules between two types of events, realized by the transfer of constraints in dynamic interactions. This serves as the physical dynamic ground of intentionality. Nevertheless, intentionality, or a representational relationship, cannot be reduced to the transfer of constraints because it is normative. In other words, a representation can be false, whereas a natural correlation cannot. A representation has a stable referent and can be about a distal event, rather than a proximal one. To fully understand intentionality, we must explain its normative aspect.

Naturalistic philosophers often turn to teleological theories to explain the normativity of representation. They conceptualize representation in terms of function. When an item or trait is selected for a particular effect, producing that effect becomes its function. In other words, it is supposed to produce that effect. For example, the function of a metal detector is to detect metal, and the function of a heart is to pump blood. Clearly, a metal detector can fail to detect metal, and a heart can malfunction, failing to pump blood. Similarly, the function of a sign is to represent something—it is supposed to represent something or carry information about an event, object, property, etc. The function of visual signals is to transmit information about the stimulus on the retinal cells to the brain; the function of the “8” waggle dance of bees is to indicate the location of a nectar source; the function of vervet monkey warning signals is to indicate the presence of predators; the function of DNA is to store and transmit genetic information; the function of the word “cat” is to represent the type of animal, a cat. Just as a tool may fail to perform its function, a sign may fail to represent what it is supposed to. Therefore, representation, as a function, has the potential to be false.

Moreover, an item may produce many effects, but not every effect is its function. For instance, a metal detector produces an electromagnetic field through a constant flow of electric current, and it also produces an effect that may harm small animals, like bugs. Detecting metal, not harming bugs, is the metal detector’s function because that is the primary reason the detector exists. Similarly, when a heart beats, it also produces sounds beyond pumping blood. Pumping blood, not making sounds, is the heart’s function because it is the reason the heart exists. Likewise,

a sign represents a distal event rather than a proximal one because propagating the constraint of the distal event is the very reason the sign exists. The waggle frequency and the angle between the crawling direction of a dancing bee and the sun are supposed to indicate the location of a nectar source, not simply that the bee is energetic and active in dancing. Indicating the location of a nectar source is the essential function of the waggle dance.

In this chapter, I discuss how a natural correlation gets its normative property to be representational relationships. [Section 4.2](#) critically reviews the leading approach to explaining the normativity of meaning in naturalized semantics, namely teleosemantics. It argues that the primary advantage of teleosemantics is that it does not require teleological notions to explain normativity, relying instead on the selected effect account of function. However, it faces the problem that what the selected effect account explains are the effects of a biological trait or item produced in the past (history), whereas what it produces are effects in the present. [Section 4.3](#) presents Peircean semiotics as an alternative approach to understanding the normativity of representation. According to this framework, interpretation plays a central role in making a correlation a representational relationship. However, interpretation is a teleological notion which needs further naturalistic explanation. In [Section 4.4](#), I support Deacon's concept of autogenesis as a plausible model for a primary interpretive system. In the last section, the concept of operational interpretation is introduced to understand interpretive processes at the unicellular level.

4.2 Teleosemantics

[Millikan \(1984; 1989b\)](#) refers to the function that an item is supposed to perform as its *proper function*. The proper function differs from its actual function and the functions that various users intend to perform on different occasions. As we have argued, a proper function can fail to be realized. For example, the proper function of a pen is to write, but someone may use it to kill. Therefore, not all effects are relevant; only the effect that an item is supposed to produce constitutes its function. When a function is proper to an item or a trait, it is not understood in a statistical sense. In other words, a proper function is not the average effect realized. For instance, the proper function of sperm is to fertilize an egg, pass on genetic information and produce the next generation. Clearly, few sperm actually fulfill this function. Representation as the proper function of signs shares this characteristic. What is properly represented by a sign is not defined by the average across idiolects. Furthermore, signs have an autonomous existence independent of individual users in a sense. What a sign represents is stabilized. Although a user may employ a sign in an improvised way to represent what it does not properly refer to, the meaning of the sign does not change based on the specific ways in which various users employ it. Accordingly, the content of a representation is determined by its proper function. A sign represents a distal event rather than a proximal one because representing the distal is its proper function.¹

By thinking of representation in terms of function, we can account for the normative aspect of intentionality. The misrepresentation problem and the

content indeterminacy problem can be solved through the function conception. However, merely considering representation as a type of function does not provide a fully naturalistic understanding because function is a teleological concept. On the one hand, function is a teleological notion and, thus, normative. Using it to explain the normativity of intentionality begs the question. We still need to explain the normativity of function. On the other hand, teleological notions are not compatible with naturalism. A core principle of naturalism is to exclude teleology from the understanding of natural phenomena, as emphasized since Descartes. In conclusion, we must explain the source of the normativity of function in a naturalistic way—that is, we must naturalize the concept of function.

For artifacts, explaining their function is straightforward. An artifact is designed to perform a certain function. For example, a coffee machine is designed to make coffee, and a hammer is designed to impact something. Some signs also acquire their proper representational function in a similar way. Humans can use any observable natural correlations to establish representational relations. For example, people use the number of tree rings as a sign of a tree's age, smoke as a sign of fire, or leaves turning yellow and red as a sign of colder weather. However, appealing to users' interpretations merely shifts the question back. Explaining the aboutness of representation through the intentional capacity of the mind still leaves us needing an explanation of the mind's intentionality. Moreover, intentionality, a hallmark of the mental (Neander 2017), is precisely what we aim to explain in the first place.

Many naturalists argue that the analogy between representation and function extends to explaining their normativity. They use the so-called etiological theory to account for function, dating back to Wright (1973). In simple terms, a function is an effect for which a trait or item was selected for in the past. That is, a trait or item exists because of the effect it produced in the past. Formally:

...for an item A to have a function F as a 'proper function', it is necessary (and close to sufficient) that ... A originated as a 'reproduction' (to give an example, as a copy, a copy of a copy) of some prior item or items that, due in part to possession of the properties reproduced, have actually performed F in the past, and A exists because (causally, historically) of this or these performances.

(Millikan 1989b, 288)

Since a proper function is defined as an effect selected in the past, this account is also referred to as the selected effects theory. For example, a metal detector is a member of a reproductive family of metal detectors with the property of detecting metal, and the detector exists because of its ability to detect metal. Similarly, a heart is a member of a reproductive family of hearts that have the property of pumping blood, and it exists because of this function. An immediate advantage of this theory is that it avoids one of the most challenging metaphysical problems teleological explanations face: The issue of future events causing present ones, a temporal converse of cause and effect. A metal detector has the proper function of

detecting metal because the same type of mechanism produced this effect in the past and was selected by designers to reproduce it.

Unlike artifacts, whose proper functions are selected by designers, there is no designer for the proper functions of biological traits. Effects as proper functions of biological traits were selected by natural selection. An effect is the proper function of a biological trait if the genotype responsible for the trait was selected to produce the effect because it contributes to the inclusive fitness of the organism's ancestors (Neander 1991a; 1991b; 1995). Roughly, inclusive fitness is measured by the number of offspring an organism reproduces. Since natural selection is a purely mechanistic process, there is no prior intention or interpretation, and the concept of function is naturalized by the selected effect theory. Representation as a type of proper function can also be naturalized in the same way. Representing a type of event is the proper function of a sign because it contributes to the inclusive fitness of the sign's users. That is, users who employed the sign to represent the event had greater inclusive fitness, and thus, the represented event was selected by natural selection. As a result, the referent of the sign was determined in this evolutionary process, without appealing to any other teleological processes.

This family of theories, based on the selected effect account of function, is known as teleosemantics. It explains the normativity of meaning. However, it is not without its challenges. Principally, there are three types of objections to the selected effect account of meaning: (1) selection history is not necessary for accounting for a function (the swampman objection); (2) selection history is not sufficient for explaining a function (the epistemological indeterminacy objection); (3) it is difficult to explain sophisticated cases of function (the sophisticated objection) (Neander 1995; Allen and Neal 2020; Schulte and Neander 2022).

The swampman is a perfect physical replica of Davidson, with no life history, emerging from a purely accidental physical event (Davidson 1987)—a thunderbolt strikes a swamp. It is difficult to deny that the swampman's heart has the function of pumping blood because it produces the same effect as Davidson's heart. Similarly, the swampman's brain states realize the same mental representations as Davidson's brain states, even though the swampman has no selection history. Therefore, selection history is not necessary for intentionality to be a function. Metaphysically, the selection history of an organism may determine a function, as we have argued. However, sometimes, it is easy to determine the function(s) of a biological trait without knowing its selection history; at other times, it is hard to understand the function(s) of a trait by appealing to its selection history. For example, people knew the heart's function was to pump blood long before they understood it was the result of natural selection. Fish living in caves without light have eyes, which were originally selected for sight, but they no longer serve that function. Similarly, the human appendix was selected for digesting leaves, but it no longer performs that function. Thus, knowing the selection history of a trait is neither necessary nor sufficient for understanding its function. It is often epistemologically difficult to determine the function of a biological trait solely by considering its selection history.

The sophisticated cases objection argues that the selected effect account is too simplistic and naïve to explain sophisticated cases of functions and meaning. On the one hand, the selected effect account relies on an overly simplistic view of evolution by natural selection, while the actual mechanisms of evolution are much more complex (Christie et al. 2022). On the other hand, a trait may have more than one function, and the function it currently performs may not be the one selected in its evolutionary history. For example, birds' wings serve both to fly and to keep warm. While bird wings were selected for flight, penguin wings no longer serve that function but instead help penguins swim.

I agree with García-Valdecasas and Deacon (2024) that a deeper challenge lies behind these three types of objections. What the selected effect account explains is the effect(s) of a biological trait or item produced in the past (history), while the effects it produces are present. It is hard to deny that the swampman's heart has the function of pumping blood because it contributes to blood circulation by circulating oxygen and nutrients, just like Davidson's heart. Cave fish's eyes no longer have a sight function because they produce no effect in the absence of light, much like the appendix in humans. They are not malfunctioning because they do not differ from other eyes. Some might argue that the eyes still have the sight function but do not perform it due to the environment's lack of cooperation (what Millikan refers to as *Normal* conditions). The case of the eyes may be controversial. However, it is clear that the human appendix no longer has a function, even though its ancestors helped herbivores digest plants. This suggests that the history of a trait does not necessarily determine its present function. More importantly, when analyzing a trait's function, we are concerned with the effect it currently produces, not the effect it produced in the past. The selected effect theory may provide a partial truth, but it is not the full story.

Furthermore, Millikan (1984, 2) clearly distinguishes the proper function of an item or trait from the functions actually performed under different conditions and those intended by specific users, emphasizing the stabilizing aspect. A proper function has an existence independent of its uses and actual performance. When the conditions are *Normal*, a proper function is reliably performed. As Millikan (1984; 1989a; 2004) repeatedly stresses, *Normal* conditions are the predominant explanatory conditions under which the function has historically been performed. *Normal* conditions are not statistically average conditions under which the function has been performed. Again, the proper function of sperm is to fertilize an egg, even though few sperm actually fulfill this function.

The concept of proper function in relation to *Normal* conditions is difficult to grasp. On the one hand, proper function and *Normal* conditions are not identical to the functions actually performed or the actual conditions; they must, therefore, be understood as dispositions or epistemological abstractions for observers. They must not be observer-dependent because the goal of the selected-effect theory is to exclude teleological elements from the concept of function. Thus, they must be dispositions. However, this understanding faces challenges. Dispositions are a metaphysical concept that can be difficult to comprehend and even appear empty. Philosophers often cite Molière's famous irony to critique the concept

of disposition: “Opium has a dormitive virtue— a disposition to make people sleep.” Dispositions explain nothing unless we frame them in non-dispositional terms. Some might argue that the concept of proper function is different because it is grounded in past history. However, the effects and conditions under which a trait or item existed in the past were necessarily different in some way. If we claim that proper function and *Normal* conditions are the common features of those past effects and conditions that contributed to the present existence of the trait or item, we would need to appeal to a statistical notion of past successful cases. This would conflict with Millikan’s idea that *Normal* conditions are not a statistical notion. Furthermore, a proper function is open-ended; it continues to evolve in the process of evolution. Similarly, we use words and sentences in different situations and ways. The actual uses of signs always vary. It seems that the selected-effect theory does not capture this aspect of function. Therefore, a satisfactory theory of representation must account for both the stabilizing and open-ended aspects.

A competing theory is Cummins’s functional analysis theory (Cummins 1975). This theory argues that a trait’s or item’s function should be analyzed in terms of its contribution to the system in which it is a component. For example, the heart pumps blood, contributing to the blood circulation system’s ability to deliver oxygen and nutrients to the organism’s tissues. This is a general theory of function, accounting for both artificial and biological functions. At first glance, this theory seems compatible with a mechanistic view, as the processes of pumping blood and its effect on circulation are mechanistic. However, this definition presupposes a teleological notion: Why do we observers determine that the heart’s contribution is to circulate oxygen and nutrients rather than hemocytes or body fluids? Why do circulating oxygen and nutrients matter? We cannot escape teleology when discussing the function in the present. We are back to where we started.

An alternative, intuitive way to explain the normative aspect of representation is by appealing to interpretation, in addition to the selected effect theory. However, as I argued in Chapter 1, interpreting an item or event as a representation of something can be arbitrary and idiosyncratic, while a representation is a reliable and stable relation between a sign and its representation. Moreover, interpretation is inherently teleological. Therefore, to explain the normativity of representation by appealing to interpretation, we must explain both the stabilized and intersubjective aspects of representation and naturalize the teleological nature of interpretation. Next, I propose a naturalized semiotic account in which interpretation plays a central role in explaining meaning, solving these difficulties (Deacon 2006; 2007; 2008; 2012a; 2017; 2021).

4.3 Semiotic understanding of information and intentionality

A tradition that has long been overlooked by naturalistic theories of meaning is Peircean semiotics.² Although semiotics is widely used in fields like literature and art, communication and media and social science, Charles Sanders Peirce’s original

aim in developing semiotics was to construct a naturalistic yet non-reductive account of the human mind, explaining its extraordinary capacity for intentionality (Short 2007, ix). This is precisely what naturalistic theories of meaning and intentionality seek to achieve. The goal of this book is to develop an alternative naturalistic understanding of information and meaning, inspired by recent developments in Peircean semiotics. In this section, I introduce the basics of Peircean semiotics and its recent expansion into biosemiotics, particularly the biosemiotic understanding of information over the past two decades.

In Peircean semiotics, briefly, to be a sign, something must represent something to some interpreter in some respect (Liszka 1996, 19). A distinctive feature of this characterization is that the concept of a sign is an irreducible triadic one—comprising the object, the sign and the interpretant—rather than a dyadic relationship between the signifier and the signified. In his classic formulation of signs, Peirce says:

A sign, or representamen, is something which stands to somebody for something in some respect or capacity. It addresses somebody, that is, creates in the mind of that person an equivalent sign, or perhaps a more developed sign. That sign which it creates I call the interpretant of the first sign. The sign stands for something, its object. It stands for that object, not in all respects, but in reference to a sort of idea, which I have sometimes called the ground of the representamen.

(CP 2.274)

A Sign, or Representamen, is a First which stands in such a genuine triadic relation to a Second, called its Object, as to be capable of determining a Third, called its Interpretant, to assume the same triadic relation to its Object in which it stands itself to the same Object. The triadic relation is genuine, that is its three members are bound together by it in a way that does not consist in any complexus of dyadic relations.

(CP 2.274)

Liszka (1996, 53) summarizes Peirce's definition of a sign with four conditions: (1) the *representative condition*, which asserts that all signs are directed toward objects, or at least purport to be; (2) the *presentative condition*, which requires the sign to represent or correlate with the object in some respect or capacity (its ground); (3) the *interpretative condition*, where the sign must determine, potentially or actually, an interpretant—understood as a sign that translates or develops the original signs; (4) the *triadic condition*, which holds that the relation between sign, object and interpretant must be triadic, forming an irreducible interrelation where each component derives its meaning from the others.

A second key distinction between Peircean semiotics and teleosemantics—a externalist approach to meaning—lies in the internal nature of Peircean semiotics.

Interpretation plays a central role in meaning-making. Deacon summarizes this central point as the central dogma of semiotics:

Any property of a physical medium can serve as a sign vehicle of any type (icon, index, or symbol) referring to any object of reference for whatever function or purpose because these properties are generated by and entirely dependent upon the form of the particular interpretive process that it is incorporated into.

(Deacon 2021, 539)

What makes something a sign is not its intrinsic properties. Rather, intrinsic properties may provide semiotic affordance. “What matters is how the relevant property is incorporated into an interpretive process, because being interpreted is what matters” (Deacon 2021, 539). This notion is similar to Millikan’s account of intentional signs, though she appeals to cooperative convention rather than interpretation to explain the normative aspect of intentionality. However, this similarity is superficial. In semiotics, there is no “infosign” as Millikan describes it; every sign involves interpretation. Thus, there are no “signs in the wild.”

Semiotics also offers a novel understanding of information, especially in its recent development, biosemiotics. Biosemiotics posits that meaning-making is an essential feature of biological phenomena. To understand life, we must explain the processes of meaning-making within biological systems. These processes are forms of information transmission, which can be understood as semiosis, or the sign process. By reconstructing Peirce’s thinking on information, we understand it as the communication of form, or habits, from an object to an interpretant through a sign, which in turn constrains the interpretant’s behavior (DE Tienne 2005; Queiroz et al. 2010; Nöth 2013; Liszka 2016). Peirce states:

...a Sign may be defined as a Medium for the communication of a Form. [...]. As a medium, the Sign is essentially in a triadic relation, to its Object which determines it, and to its Interpretant which it determines. [...]. That which is communicated from the Object through the Sign to the Interpretant is a Form; that is to say, it is nothing like an existent, but is a power, is the fact that something would happen under certain conditions.

(EP2, 544, 22)

This understanding of information revives its etymological meaning—the action of bringing form into material (Peters 1988; Capurro 2009). It also explains the unidirectional, dependent relationship between object, sign and interpretant, where the object determines the sign, which then determines the interpretant—not vice versa. The central questions in information theory thus become: What is form? Where does form exist in nature? And how does form acquire its meaning within semiotics?

These questions are particularly important for understanding biological information. Hoffmeyer and Emmeche (1991) combine semiotics with second-order

cybernetics to explain this. They argue that biological information emerges from the unique ability of living organisms to respond to selected differences in their environment to maintain themselves. Information is produced in this process, and to understand it, we must comprehend this distinctive ability. “For a system to be living, it must create itself, i.e., it must contain the distinctions necessary for its own identification as a system” (Hoffmeyer and Emmeche, 1991, 126). This self-referential feature is essential to life. At its core, this is code-duality: The ability of a system to represent itself in two different codes, one digital and one analog, and to transform one into the other. In the transformation of analog coding to digital, meaning is made, and a physical form gains significance and relevance for the organism. Conversely, when the digital is transformed back into the analog, the organism’s intention has a real, physical effect on the world. Semiosis, realized through code-duality, bridges the living and the lifeless. According to this view, the normativity of intentionality originates from code-duality.

The next question is: How can a form be about another? Instead of viewing the referential relationship of aboutness as monotypic, Peirce classifies three types of basic signs in terms of their referential relations: Icon, index and symbol. An icon correlates with its object through likeness or similarity; an index correlates with its object through contiguity, whether causal or spatial-temporal; and a symbol correlates with its object through a conventional, habitual relationship. Furthermore, these three types of signs are hierarchically nested: A symbol grounds on an index, which in turn grounds on an icon (Deacon 1997). This classification may offer a radically new perspective on classic problems in linguistic philosophy, such as the problem of sense and reference, the misrepresentation/fiction problem and the symbol grounding problem.

Despite Peircean semiotics being a sophisticated theory of intentionality, most philosophers overlook it, and there are strong reasons for this. First, appealing to interpretation to explain the normativity of intentionality seems to postpone the real question without answering it fully. Within the semiotic framework, the real problem appears to be explaining the competence of interpretation, something semiotics assumes rather than addresses. Second, as Babieri (2007; 2019) argues, appealing to interpretation may not align with naturalism. While Peircean semiotics provides a descriptive explanation of semiosis, the mechanisms behind how semiosis is possible are left unexplained. Clearly, semiosis is inherently teleological in the Peircean framework, and we need a mechanistic account of the intrinsic purpose of interpretive systems, if such purpose exists. Third, because interpretation is individually subjective, it is difficult to explain the conventionality, intersubjectivity, or even objectivity of meaning. Fourth, the triadic model of signs remains somewhat obscure, making it challenging to apply it to solve the problems raised by naturalistic theories of meaning, such as the reference-class problem and the strength problem.

If semiotics is to contribute to the naturalization of intentionality, an account of naturalizing interpretation and types of semiosis is needed. In this book, I will argue that the emergent dynamics theory and biosemiotics theory proposed by Deacon (2006; 2012a; 2018) may resolve several of these problems. In the next

section, I will explain what a minimal interpretive system is, through the toy model proposed by Deacon.

4.4 Autogenesis and the origin of normativity

In his recent paper “How Molecules Became Signs?,” Deacon argues that to address the question, “What sort of process is necessary and sufficient to treat a molecule as a sign?” we must focus on the interpretive system (Deacon 2021). For him, explaining the interpretive system means explaining what kind of system a living system is. This implies that life itself is sufficient for interpretation. Two questions need to be answered in this context. First, what is the simplest system that is alive? Second, is the system sufficient for interpretation? To answer these two questions, I will employ Deacon’s thought experiment of the simplest teleological system, autogen (Deacon 2006; 2012a; 2020).

Understanding the nature and origin of life is one of the most challenging problems in life sciences. Broadly, two approaches to this problem exist: The Darwinian and the organismic (Godfrey-Smith 2016).³ The Darwinian approach argues that the nature of life lies in the mechanism of replication, where molecular information is transcribed to facilitate self-replication and self-production. In contrast, the organismic approach emphasizes the self-producing and self-maintaining characteristics of life (to name a few, Maturana & Varela 1972/1980; Eigen and Schuster 1979; Gánti 1979/2003; Varela 1997, Kauffman 1993; 2000; Rasmussen and Bedau et al. 2009). From this perspective, the nature of life is defined by self-generation and self-persistence, achieved through resistance to thermodynamic equilibrium via metabolism. While the Darwinian approach captures the informational aspect of life, it neglects the physical aspect. On the other hand, the organismic approach, with its high requirements for a living system, cannot realistically serve as a primary model for the origin of life, nor does it account for the informational aspects of life. While it is debatable what the essential features of life are, it is reasonable to assume that a satisfactory account should encompass the concerns of both the Darwinian and organismic approaches while avoiding their shortcomings. Deacon’s toy model of autogenesis may meet these requirements.

Generally, three elements are crucial for a system to be considered living: Metabolism, genes and containment⁴ (Gánti 1979; Deacon 2006; Rasmussen, Bedau 2009 et al.). Metabolism defines the self-producing nature of life. A living system produces the elements that constitute it through metabolism, enabling self-repair, self-reconstitution and self-persistence, thus maintaining self-generation and self-maintenance. Containment is the boundary that separates the internal from the external in an organism. At the same time, it serves as the interface through which the organism interacts with its environment. Genes are ubiquitous in living systems and essential for self-replication. A satisfactory account of life’s nature must explain these elements. I will now argue that self-replication, metabolism and containment can be spontaneously realized in the thought experiment of autogenesis.

The simplest autogen is realized by two reciprocal self-organizing processes: Reciprocal catalysis and self-assembly (Deacon 2021). A reciprocal catalysis

involves at least two catalytic reactions, where the product of each reaction catalyzes the other. Given a sufficient number of substrate molecules, this reciprocal catalysis persists and can expand into a network of cyclic reaction chains. Self-assembly, on the other hand, is a type of molecular aggregation process in which energetically favored molecular components spontaneously aggregate in localized regions. This process results in the formation of large, closed, regular structures, such as polyhedral or tubular capsids.

These two processes complement each other. When coupled, each provides critical boundary conditions for the other. As the reciprocal catalysis network grows, it becomes increasingly fragile and prone to diffusion. However, when the self-assembled capsid forms containment around this network, it isolates it from external interference, effectively preventing catalyst dissipation. In turn, maintaining the capsid containment requires a high local concentration of specific molecular components, which the reciprocal catalysis network can produce as a byproduct. As these processes integrate, the previously externally imposed critical boundary conditions become intrinsic to the new whole. Deacon refers to this integrated system as an “autogen.”

When reciprocal catalysis is initiated, provided appropriate substrate molecules are present, it produces a single species of component molecules that tend to self-assemble into a closed capsid containment. This structure then encapsulates the reciprocal catalysis, forming an autogen. As the substrates inside the containment are consumed, reciprocal catalysis ceases, halting the production of component molecules. The containment becomes vulnerable to damage without a continuous supply of components, eventually opening up again. If appropriate substrates are nearby, reciprocal catalysis can restart. With a new supply of components, the containment self-repairs. If the damage is extensive enough to split the containment into separate parts, these independent fragments may restart separately when required substrates in the surrounding environment are rich, forming different autonomous autogenes. Thus, autogenes have the capacity for self-reproduction.

As we can see, the account of autogenesis aligns with the organismic understanding of life. Nevertheless, this account differs from other theories of life in important ways. First, some theories require a system to have a persistent metabolism and a semipermeable membrane to be considered living. However, an autogen does not possess these properties. When the capsid structure forms the containment, autocatalysis, which is analogous to metabolism in unicellular organisms, stops. In other words, there is no persistent “metabolism” in the autogen. Second, existing theories of life argue that a semipermeable lipid membrane is necessary for a living system. Such a membrane serves not only as a boundary separating the system from its environment but also as an interface through which the system interacts with its surroundings. The containment in an autogen is not semipermeable. When it is closed, it completely separates the inside from the outside. However, it also serves as an interface. When the containment is broken, the autogen interacts with its environment.

Furthermore, autogens are evolvable. When an autogen exhausts its reactions and reopens to its environment, unanticipated components from the surroundings

may be incorporated into its containment. Some of these may disrupt the internal reciprocal catalysis, leading to the autogen's destruction, while others may enhance the efficiency of catalysis, resulting in autogens that replicate more efficiently. Some newly incorporated components may neither contribute nor disrupt the original processes. In the case of more efficient autogens, which reproduce descendants in a more effective manner, their fitness surpasses that of the original autogen. As these descendants inherit the same structural information, they gradually take over the population. In other words, evolution by natural selection occurs.

Some may argue that although the autogen is an intriguing toy model, it is not a living system because it does not capture the informational features of life. An independent molecular template, responsible for information replication, is ubiquitous in living systems. Even viruses possess such templates, realized by RNA instead of DNA. These templates encode and store genetic information that can be used to replicate another system with the same replication capacity. According to [von Neumann \(1966\)](#), a system capable of replication can use the information encoded in the template to construct a system with the same structure and function, meaning it encodes the same information and has the same capacity for replication. Von Neumann argues that for a system to have the capacity for replication, it must have three components: A description of itself (information about itself encoded in the template), a universal constructor mechanism that reads the template's information and constructs the system, and a universal copy mechanism that copies the information. The molecular template thus plays a central role in replication, which is essential for life according to Darwinian theories of biological individuality. There is no corresponding part in the autogen that performs the replication of information.

It is true that there is no independent molecular template encoding genetic information like in existing living systems. However, this does not imply that the autogen lacks a replication function. In fact, the structure and dynamic constraints of the autogen realize the functions of description and construction simultaneously. Both the reciprocal catalysis network and the containment have structures that encode structural information—information about the location and relationships between components. This structural information constrains the available components and the ways they can be constructed, thus serving as a description of the autogen itself. When conditions are favorable, the reciprocal catalysis network generates new components, which then construct the containment according to the constraints of the structural information, leading to growth. Therefore, it performs the function of a universal constructor.

Now, imagine that an autogen is severely damaged into several fragments, losing its capacity for self-repair. Despite this, each fragment retains some part of the structure containing the autogen's structural information. When the local environment provides enough reactants to restart the reciprocal catalysis, each fragment has the opportunity to regenerate into a new autogen with the same structural information. These new autogenes will have the same structure and function as the original. Thus, we can say that the autogen performs replication, and the autogen model does capture the informational aspect of life.

Some may argue that, although the structure of an autogen contains structural information, as I argued in [Chapter 3](#), structural information or constraint is not yet information. It seems that I am smuggling in the very thing the account of autogenesis seeks to explain—interpretation—in an attempt to understand the replication of the autogen. Therefore, my argument seems to fall into a vicious circle. To resolve this, I must demonstrate that an autogen is a purposive system with the capacity for interpretation. Deacon has argued that autogenes exhibit normativity and interpretive competence ([Deacon 2021](#), 546). In addition to normativity and interpretive competence, Deacon asserts that autogenes possess three other holistic properties: Individuation, autonomy and recursive self-maintenance ([Deacon 2021](#), 546).

Through its containment, the autogen physically and functionally distinguishes itself from its surrounding environment. It achieves autonomy because “it intrinsically embodies and maintains its own boundary conditions via component processes that reciprocally produce the external boundary conditions for each other.” Furthermore, “it repairs and replicates the critical boundary conditions required to repair and replicate these same critical boundary conditions” ([Deacon 2021](#), 546), thereby achieving recursive self-maintenance. The autogen can be seen as an intrinsically teleological system in a Kantian sense: Every part exists through all other parts, and in turn, every part is for the other parts and for the whole ([Weber and Varela 2002](#); [García-Valdecasas 2022](#)). Every process constituting the autogen exists for the sake of its continued existence. Therefore, the autogen is a purposive system.

With the emergence of intrinsic purpose, normativity follows. As each process acts for the persistence of the whole, it may fail. For example, damage to the containment or unintended involvement of other molecules in reciprocal catalysis can disrupt the self-assembly of component molecules, leading to the failure of the autogen. By clarifying normativity, we can better understand the interpretive competence of the autogen, which can be understood through the concept of code-duality proposed by [Hoffmeyer and Emmeche \(1991\)](#). Next, I will argue why autogenesis is an interpretive process and how structural information or constraint transforms into signs with meaning through code-duality.

4.5 Genetic information, code-duality and operational interpretation

To understand why autogenesis is an interpretive process, let’s revisit a classical problem in the philosophy of biology: Whether genetic information is semantic. Genetic information plays a central role in molecular biology, and as a result, the conception of information dominates the field. Terms such as transcription, translation, transduction, code, redundancy, synonymy, messenger, editing, proofreading and library are all technical terms in biology. The informational implications of these terms have been heavily debated between biologists and philosophers. Some argue that information-related concepts in biology are truly semantic and representational, while others disagree. Those who reject the semantic claim argue that

biologists merely use these terms metaphorically, underpinned by a mechanistic understanding of life (for example, Sterelny et al. 1996; Sarkar 2000; Griffith 2001; Levy 2011). Supporters of the semantic claim agree that genetic information is indeed semantic, but they understand the concept of information in different ways. Some view it through a teleosemantic lens (Maynard Smith 2000; Shea 2007), while others consider it within the framework of information transmission, where interpretation plays a central role (Jablonka 2002; Bergstrom and Rosvall 2009).

I argue that information does play a substantial role in biology, but not in the sense of teleosemantics. While some appeal to the interpretation systems of receivers or observers, they do not provide a clear account of interpretation or explain why processes such as translation, transcription, transduction and replication of genetic information are interpretive in nature. In this section, I argue that a mechanistic understanding of genetic information is insufficient and provide an account of operational interpretation to better understand it. Let's begin with the central dogma of molecular biology.

The central dogma of molecular biology, which plays a pivotal role in the field, introduces the concept of informational processes in biology to a large extent. It suggests that the flow of genetic information during protein synthesis is a purely mechanistic process.

The Central Dogma. This states that once “information” has passed into protein it cannot get out again. In more detail, the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible. Information here means the precise determination of sequence, either of bases in the nucleic acid or of amino acid residues in the protein.

(Crick 1958: 153)

According to the principle, sequence information is initially coded by the sequential structure of DNA. During replication, this information is transcribed and encoded by messenger RNA (mRNA) through template matching (transcription). The mRNA, carrying genetic information, then forms a complex with ribosomes and transfer RNA (tRNA), serving as a template for protein synthesis. This process ensures that proteins fold into the correct structures based on the genetic information represented by the mRNA.

Although there are transfers of genetic information that fall outside the scope of the central dogma, it still holds significant importance in molecular biology. However, my concern here is not whether the dogma can be upheld, but whether it is purely mechanistic. As we can see, several concepts with obvious teleological connotations, such as code, information, transcription, translation and correctness, are integral to this process. How can these concepts be reconciled with a mechanistic explanation?

A common response to this issue suggests that these concepts are often used metaphorically, with the belief that they can ultimately be explained in physical terms. For instance, biologists use the terms “code” and “information” in a highly

restricted sense, as opposed to the richer, more semantic notions they typically imply (Sarkar 1996; Godfrey-Smith 2000; [Griffiths 2001](#)).

...there is one kind of informational or semantic property that genes and only genes have: coding for the amino acid sequences of protein molecules. But this relation ‘reaches’ only as far as the amino acid sequence. It does not vindicate the idea that genes code for whole-organism phenotypes, let alone provide a basis for the wholesale use of informational or semantic language in biology.

([Godfrey-Smith and Sterelny 2016](#))

Moreover, some argue that the genetic code exhibits arbitrary characteristics because many other mappings between DNA base triplets and amino acids are biologically possible. Yet, the perceived arbitrariness of the genetic code is superficial, stemming from gaps in our understanding of the complex connections between DNA base triplets and amino acids. As [Godfrey-Smith and Sterelny \(2016\)](#) suggest, the very notion of arbitrariness in this context proves elusive.

In their groundbreaking work, [Hoffmeyer and Emmeche \(1991\)](#) argue that the flow of genetic information in protein synthesis should be understood as interpretive acts, or semiosis. They propose that the concept of information needs to be redefined through Peircean semiotics. If genetic information is understood simply as the determination of sequence, it becomes indistinguishable from other forms of organization in nature. What distinguishes genetic information from other biological information during ontogenetic development is its semiotic nature: it functions as signs that stand for something within interpreting systems. While genetic sequences are relatively independent and stable within cells, they should not be understood in isolation. Instead, sequence information should be considered within the broader context of ontogenetic development and reproduction, in which it is interpreted. Hoffmeyer and Emmeche describe semiosis at the organism level as follows:

In the epigenetic process, DNA, i.e. the genome, may be seen as just one fragment of an evolutionary stream of signs passed down through the generations. The interpretant selecting such signs among the myriad of internal cytoplasmic differences is the fertilized egg, the zygote. The fertilized egg, the real ‘person’ of biology, can make sense of the elaborate message contained in the DNA, using it to master the epigenetic process, i.e. the construction of the phenotype, the actual organism.

([Hoffmeyer and Emmeche 1991](#), 143)

In this semiosis, “[s]till, it is the zygote which is the subject in the process: It initiates the deciphering of the DNA-message and becomes gradually changed to the embryo in response to the interpretation” ([Hoffmeyer and Emmeche 1991](#), 144). Unlike molecular biology, which interprets ontogenetic development as biochemical processes guided by genetic information, biosemiotics view it as a form

of semiosis. This perspective emphasizes that the process can only occur within an interpreting system, such as the zygote. Because the zygote acts as an interpreting system, the normativity of the genetic code (information) is no longer a mystery. However, similar to other theories that appeal to interpretation, an account of interpretation that can be employed to understand basic semiosis like genetic information is still needed. We must understand in what sense these processes are interpretive.

Hoffmeyer and Emmeche (1991) employ code-duality to elucidate how semiosis functions in protein synthesis. Code-duality refers to the recursive transmission of information through the interactions between digital and analog codes (Hoffmeyer 2008: 80). Digital codes are those in which the components are discrete tokens connected by arbitrary relations, while analog codes rely on similarities in spatio-temporal, part-to-whole or causal continuity. Digital codes are used for memory storage, while analog codes facilitate decoding and instructing physical realization. According to the concept of code-duality, genetic codes function as digital codes, while the processes of transcription and translation operate as analog codes. Similarly, interpretation (semiosis) within autogens also operates through code-duality, albeit in a distinct manner.

Unlike the code-duality observed in protein synthesis, autogens do not exhibit physical differentiation. As previously discussed, the two self-organizing processes in autogens mutually provide critical boundary conditions. These boundary conditions serve as constraints, limiting the possible realizations of the two processes and ensuring the autogen as a whole. Through self-maintenance and self-reproduction, the autogen preserves these constraints. “This preservation of constraints both provides a record and a source of instruction for organizing the work required to preserve this same capacity” (Deacon 2021, 545). This means that the constraints embodied in the interplay of these two coupled processes serve dual roles, akin to both digital and analog codes.

The constraints in autogens are digital because they are arbitrary. If reciprocal catalysis and self-assembly were uncoupled, each could proceed in numerous other ways. However, when coupled, they mutually constrain each other to operate within a much more limited set of possibilities, thereby sustaining the autogen as a whole. Moreover, the dynamic structure of the autogen re-represents its own boundary conditions. Essentially, these constraints store the sequence information of the autogen itself. When an autogen reproduces another, these constraints are preserved and intrinsically represent and reproduce the critical boundary conditions necessary for the autogen’s existence (Deacon 2021, 546). On the other hand, these constraints are analog because they are realized through spatio-temporal (self-assembly) and causal continuities (reciprocal catalysis). We can describe the dynamic structure of an autogen as digitally coding information about its boundary conditions. When damaged, the remaining parts initiate an analogous decoding process to self-repair or self-reproduce, given adequate environmental conditions.

This self-referential code-duality realizes the most primitive form of interpretation, implementing the primitive semiosis. The dynamic structure of an autogen represents itself (the representative condition); it re-presents the critical boundary

conditions of the autogen (the presentative condition). When an autogen is damaged, its integrity is disrupted and self-repair or self-reproduction is initiated, generating a new autogen(s) as the interpretant (the interpretative condition). The relations between the autogen (object), dynamic structure (sign) and regenerated autogen (interpretant) constitute genuine triadic relations that cannot be reduced to any dyadic relation between any two of them (the triadic condition). Therefore, autogenesis meets the four formal conditions of signs given by Peirce. The next question is what kind of semiosis autogenesis represents. Since it involves the regeneration of self/non-self distinction, it is considered an iconic semiosis (Deacon 2021, 546).

This primitive form of semiosis possesses several distinguishing properties from semiosis in the general sense. First, its reference is itself, making it inherently self-referential. Second, the autogen (object), dynamic structure encoding constraints (sign) and regenerated autogen (interpretant) exist not as separate entities but as a persistent unity: Autogenesis. Third, it provides a diachronic instantiation of semiosis, contrasting with synchronic forms. The primitive semiosis manifests in the dynamics of autogenesis, which self-referentially structures itself in irreversible time. Fourth, the dynamics of the autogen are physically determined and thus mechanistic, aligning with modern scientific methods and experimental feasibility. Due to these distinctive characteristics, I propose calling this primitive form of interpretation “operational interpretation.” Other forms of interpretation (semioses) can be differentiated from operational interpretation, as Deacon (2021) has argued. With the operational interpretation concept in mind, let’s return to the normativity problems of representation.

4.6 From constraint to information, and to semiosis

According to the notion of operational interpretation, when incorporated into autogenesis, a constraint imposed by reciprocal catalysis and self-assembly becomes meaningful. The structural information embodied by the constraint becomes referential, referring to the autogen itself. It is also normative with respect to its contribution to the self-maintaining and self-producing autogenesis. When a process fails in its contribution, it malfunctions accordingly. The constraint, or structural information embodied in the process, is no longer about itself, and as a result, it carries misinformation. Furthermore, the constraint is propagated to newly generated parts in self-repair and to newly formed autogens in reproduction. We may say that the information is transmitted from old parts to the new, from parent to daughter. Nevertheless, this self-referentially iconic reference pertains to itself. It is not the typical referential relationship in our minds, where what we usually take to be a representation is about something else. Therefore, this simplest form of autopoiesis is too simple to investigate the normativity of representation, to understand the relationship between history and the present referential function of a representation, and to resolve the tension between the stabilizing and open-ended aspects of the function. A more complex model is needed.

Now imagine a slightly different autogen with a sensitive surface (Deacon 2012a; 2021). Its capsule has molecular features to which the substrate tends to bind. As the number of binding substrates increases, the structural stability weakens, and thus, the bonding points become more fragile. As a result, the autogen tends to open its containment more frequently in a substrate-rich environment. When the containment is closed, the autogen is stable. However, if there is substrate concentration in the surroundings, the containment becomes more likely to be disrupted. When the containment opens up, it is more probable for the autogen to replicate itself, since there are rich substrates nearby. With the replication of more and more autogens, local substrate would be quickly used up. The containments of autogens close up, restoring their stability. Comparatively, the sensitive autogen has an adaptive advantage over the original generation. The original generation is insensitive to its environmental conditions and thus retains its stability even when there are abundant substrates around. Consequently, the autogen is less likely to replicate as the evolved ones do. That is, the fitness of the sensitive autogen is much higher than that of the original.

The adaptive advantage of the sensitive autogen comes from its novel semiotic competence: Indexical representation. Semiotically, an indexical representation is spatio-temporally or causally continuous with its referent. A left-turn arrow traffic sign is an index. The direction of the arrow is continuous with the direction it indicates. A driver interprets the arrow sign as indicating the direction they should go forward in the lane at the junction. In the case of the sensitive autogen, the binding of the substrate is continuous with the suitability of the environment. The bonding point that binds the substrate to the surface is an index of the environment's suitability. The bonding point is a sign, and the substrate is the indicated object. Their continuity is operationally interpreted by the sensitive autogen. That is, with increasing binding of the substrates, the autogen's containment is disrupted, and the substrate gains access to the interior, engaging in reciprocal catalysis. As catalysis accelerates, the autogen self-repairs or produces replicated autogens. These outcomes are the interpretants. In terms of the central dogma of semiotics, the binding of the substrates, disrupting the containment, and containing the substrates all contribute to the self-repairing, self-maintaining, self-producing and self-replicating functions of an interpretive system: The sensitive autogen. This makes the entire process more than merely a chemical reaction—it becomes semiosis.

In this thought experiment, more specific details deserve further articulation. First, without being incorporated into the autogenesis, the properties of the substrates are merely semiotic affordances—possibilities available in the environment for an interpretive system to use (Chemero 2003; Heras-Escribano 2020; Deacon 2021, 539). Furthermore, not all properties of the substrates, but only those that are relevant to the self-maintaining and self-producing processes, are involved in semiosis. In other words, the substrates ground the sign (the bonding point on the autogen's surface) in the sense that the sign only presents the properties of the substrates that matter to the autogen, rather than all of its properties. This is the presentational condition for being a sign.

Second, in Peircean semiotic terms, the presence of the substrates in the environment is the dynamic object represented by the sign, while the general suitability of the environment is the immediate object (Deacon 2012a, 443). The general suitability of the environment for the autogenesis' self-production and self-reproduction should be understood in terms of negative logic. When conditions that prevent the self-production of the autogen are absent in the environment, the environment is considered suitable. Therefore, suitability is a negative property for the autogen. As a result, it is multiply realized. It is realized by the substrates in the present. Hence, the immediate object is grounded by, but not reduced to, the dynamic object. The part properties of the dynamic object that contribute to the autogenesis also contribute to the environment's suitability. Here lies the problem of content determination: What does the bonding point indicate, the substrate (dynamic object) or the general suitability of the environment (immediate object)? I would argue that this is a misguided question. Positively, the presence of the dynamic object constitutes part of the immediate object. Negatively, the selection of a specific ground for representation is a function of, and therefore constrained by, the general suitability for the self-production of the autogen, which is defined according to the autogenesis. This implies that what a sign represents is determined by its contribution to the self-maintaining and self-producing processes of interpretive systems. Therefore, while we may classify immediate and dynamic objects for some epistemological purposes, they are not significant in the operationally interpretative process.

Third, in Peircean semiotic terms, the resulting repaired and descended autogens are interpretants. However, some may argue that, although the sensitive autogen reopens its containment by increasing substrate binding and begins replicating when there are abundant substrates, the presence of the substrates and their contribution to the autogenesis may be accidental and idiosyncratic. Therefore, it is difficult to claim that the proper function of the bonding point is to represent the general suitability of the environment. Teleosemantics appeals to evolutionary history through natural selection to explain the stabilizing aspect of representation's normativity. However, as argued in Section 4.2 a representation, as a stabilizing function, is about the past according to the etiological account of function, while a living system uses representation to address present conditions.

Peircean semiotics distinguishes between dynamic and final interpretants. The dynamic interpretant is the direct or actual effect produced by a sign upon an interpretive system, while the final interpretant is any rule-like or law-like effect a sign has upon any interpretive system (Liszka 1996, 26–27). In our thought experiment, the repaired autogen and its replicants are dynamic interpretants, while the perpetuation of this interpretive habit through the persistence of the autogen lineage (Deacon 2012a, 443) is the final interpretant. But how is the final interpretant possible? What is the relationship between the final and dynamic interpretants? I also appeal to evolutionary explanations by natural selection, but in a negative way.

The negative understanding of evolution suggests that evolution does not determine or produce a function, as the selected effect theory claims, but rather imposes constraints that scaffold how a function is performed. A basic recipe for natural selection includes natural variation, differential reproduction and inheritance

(Lewontin 1970; Godfrey-Smith 2007; Skyrms, 2010a, 50). In brief, there are variants in biological traits among members of a population, different variants have different reproductive success (fitness), and the variations are inheritable. With these three factors, evolution proceeds. A variant with higher fitness will prevail in the population, while those with lower fitness will be gradually eliminated. Natural selection is a passive mechanism. It does not generate novelty. It is indifferent to the mechanisms of variation and inheritance. It merely eliminates those that do not fit the environment. What remains in the population is the result of eliminating other variants. In other words, natural selection works in a negative manner. It increases the likelihood of survival and reproduction for variants that passing through constraints, while eliminating possibilities for variants that do not fit. This implies that natural selection does not select for specific traits but narrows the range of possibilities for biological traits.

Some may argue that the stabilizing function of a trait is very specific, and thus, it is hard to say that the trait is not selected for its effect. For example, human eyes are specifically adapted for vision. If eyes are not selected for vision but are the result of constraints, how could they be so sophisticated for enabling vision? I would argue that this is another version of Paley's watch argument. Paley argues that if we find a watch on the beach, it must have been designed by someone because it is too sophisticated to have formed spontaneously. Similarly, eyes are too sophisticated to be merely the result of narrowing the space of possibilities. Even if the range of possibilities is constrained, it remains vast. Therefore, eyes must have been selected for the effect of enabling vision.

Like Paley, who ignores the timescale on which natural selection operates (geological time) and treats natural selection as random, this argument overlooks the scaffolding effect in evolution. Natural selection may have been random when life first originated. However, once evolution begins, it shapes constraints upon constraints and inevitably generates increasingly complex forms of dependency (scaffolding) on those constraints (Deacon 2008, 188). In other words, if a species evolves, it must begin with the constraints it has embodied. Deacon (2012a, 318) calls this effect the "ratchet effect." That is, evolution can only proceed in a direction constrained by the past. In other words, evolution must begin with constraints shaped by history; it cannot ignore the past. Constraints formed in the past scaffold present and future evolution. Given enough time, it is entirely possible to evolve a trait that seems to have been specifically selected for the effect it produces, from the perspective of an observer.

An alternative advantage of this negative account of evolution is that it views evolution as an open-ended process. If a trait's properties are not constrained by existing factors, they can be inherited by its descendants. A trait typically has many properties. Some properties may improve fitness under certain conditions, while others improve fitness under different conditions. For example, the feathers of birds initially evolved for warmth may later serve the function of flight. These phenomena are called exaptation and preadaptation by biologists (Gould and Vrba 1982). In summary, the negative account of evolution explains the tension between the stabilizing and open-ended aspects of a biological function.

With this negative understanding in mind, let's revisit the relationship between the dynamic and final interpretants. The constraints shaped by evolutionary history form the foundation of the final interpretant. The final interpretant does not determine what a sign represents. Instead, it constrains the possible interpretations of the sign—what a dynamic interpretant can be. We may say the distinction between dynamic and final interpretants is similar to Millikan's distinction between stabilizing function and actual function. However, the similarity is superficial. Unlike the stabilizing function, which teleosemantics treats as the function a sign was selected for in the past, the final interpretant is negatively shaped by the past and serves as a constraint on the actual performance of a sign. Unlike the function that is actually performed, which includes both successful and failed performances, the performance of the dynamic interpretant is constrained by the final interpretant. A sign may fail to produce a dynamic interpretant if it does not contribute to the self-maintenance and self-production of the interpretive system it is involved in.

Now we have a naturalistic semiotic account of representation. It proposes that to understand representation, we must explain the interpretive system—a living system that performs interpretation—and evolution, which shapes constraints on the interpretive system. In other words, the representational function of a sign should be understood in two dimensions: Evolution and interpretation. As [Deacon \(2008, 188\)](#) states:

The evolved mechanism constrains the dynamics of possible interpretation, but doesn't determine it. Each moment of interpretation is in some way supportive or disruptive of the self-maintenance of this dynamical trend. This means that not only is there an historical origin for the normative property of this interpretive process, there is also an ahistorical and immediately efficacious normative property as well.

([Deacon 2008, 188](#))

This naturalistic semiotic account has several advantages over teleosemantics. First, it explains the normativity of representation rather than simply explaining it away, as teleosemantics does. Teleosemantics accounts for the normativity of representation by appealing to evolution by natural selection. On one hand, evolution by natural selection is a completely mechanistic process. Since representation is a product of evolution, its normativity appears superficial. On the other hand, evolution by natural selection assumes, rather than explains, a living system. The autogenesis account, however, explains the fundamentals of living systems and understands normativity in relation to the self-maintenance of these systems.

The misinformation problem and the content determination problem are thus resolved. If a representation is about something that is not constrained by its history and consequently fails to perform the evolved function related to self-maintenance, it malfunctions and carries misinformation. Which aspect of the represented object (the immediate object in Peircean semiotic terms) grounds a representation is

determined by the representation's history and by the properties that currently contribute to the self-maintenance of an interpretive system. Furthermore, this account avoids falling into the idiosyncratic trap of appealing to individual interpretation to explain normativity because interpretation is constrained and thus stabilized by its historical context.

Second, this naturalistic semiotic account explains how a representation, stabilized by past conditions, can still be about present conditions. According to teleosemantics, the stabilizing function of a representation is determined by its past. Consequently, what the representation should refer to is the past condition. However, a representation is typically used to refer to present conditions. The proposed account argues that when a representation is involved in an interpretive system, it refers to present conditions, though constrained by its past.

Third, this account resolves the tension between the open-endedness and stabilization of representation. Teleosemantics leaves no room for the open-endedness of representation, while living systems continually use representations in more or less different ways under varying conditions. By understanding evolution from the perspective of negative logic, the naturalistic semiotic account explains the stabilizing aspect of representation while also allowing for the possibilities of evolution in how the representation can develop.

With this naturalistic semiotic account in place, we provide a constructive theory of structural, referential and normative information. Structural information is naturalized as constraints, and referential information is grounded in constraint propagation through dynamic interactions between different processes. A constraint propagation becomes meaningful only when it is involved in an interpretive system (the central dogma of semiotics). As Deacon says,

Function and representation are made possible by the way living processes are intrinsically organized around absent and extrinsic factors, and the Darwinian process inevitably generates increasingly convoluted forms of dependency on absence. Information is a relational property that emerges from nested layers of constraint: constraints of signal probability (Shannon), constraints of the dynamics of signal generation (Boltzmann), and the constraints required for self-maintaining far-from-equilibrium end-directed dynamics (Darwin).

(Deacon, 188)

This naturalistic account of information and representation explains only very simple and ideal representations. However, representation can be far more complex and convoluted, especially in language. For example, fiction is pervasive in linguistic expression, with instances like "Harry Potter" or "the French King in 2025." Simply explaining fiction as misrepresentation or pretending representation is unsatisfactory and difficult to accept. A complete theory should also be able to explain these more complex cases of representation. This will be the task of the next chapter.

Notes

- 1 When I say the content of a represented is determined, I say it in a metaphysical sense than epistemological. Metaphysically, an item or a trait has one or several proper functions. It does not mean that we can discern what is the proper function of a trait. This is the epistemological difficulty Dretske's bacteria case tries to capture. However, what I concern here are the metaphysical questions, so I will leave the epistemological ones aside though not irrelevant.
- 2 Some may argue that this claim is not fair to Millikan who clearly says that her biosemantics develops from Peirce's semiotics ([Millikan 1984; 2004; 2017](#)). However, I will argue that her reformed theory of signs is radically different from Peirce's semiotics.
- 3 I cannot do a fair treatment of the discussion on the nature of life due to the limited space here. People who have interest in the topic may read [Bedau \(2008\)](#).
- 4 This is a working hypothesis. It does not deny that there are boundary cases which do not have one or two of the elements. For example, a seed of sleep almost has no metabolism. It is hard to define the boundaries or containments of many organisms, like plants, bacteria and fungi.

5 Reference

5.1 Introduction

A naïve view of representation is that it simply refers to or represents something. For example, a portrait represents a person; the waggle dance of a bee refers to the location of nectar; vervet monkeys' alarm calls represent the appearances of corresponding predators; the left-direction arrow on the road indicates the direction of a lane; and the sentence "Today is cold" represents the state that today is cold. In naturalistic accounts of meaning, it is often said that the proper function of a representation is to represent. That is, a representation is designed by either users or evolution to serve this function. If a representation fails to represent, it is considered a misrepresentation, and people argue that the truth conditions under which it could successfully represent have not been met. This naïve idea applies to all types of representation, from various biochemical signals at the subpersonal level to animal signals and, of course, to language.

According to Wittgenstein's picture theory of meaning, language operates as a system of representation in which words and sentences represent states of affairs in the world (Wittgenstein 1921/2010). Language as a whole represents the world. It shares the same logical form as the reality it represents. In other words, the structural relationships between components of language mirror the relationships between objects in reality. However, linguistic referential relationships differ from others. First, unlike animal signals, which directly represent objects, properties or events, a word or phrase (other than proper names) cannot refer directly to any object unless it is embedded in a sentence or specific context. For instance, while the word "dog" refers to the animal, it does not refer to any specific dog unless it is a part of a sentence that provides further context. Second, language can refer to something general and abstract, such as almost all logic and mathematical concepts like "triangle," "function" or "deduction." Third, language can be used to refer to things that do not exist, namely fiction. Examples include "Harry Potter" and "Middle Earth." Fourth, people often use linguistic expressions in ways that do not align with their literal meanings, such as irony or metaphor. Understanding how these phenomena of reference and meaning work is a crucial task for the philosophy of language, philosophy of mind and philosophy of cognitive science.

Linguistic philosophers have made significant contributions, but they have reached few agreements. This chapter aims to develop an alternative approach to the problems of reference. Following Deacon's work (1997; 2012b; 2018), I argue that traditional theories fail to explain linguistic reference because they overlook the complex semiotic infrastructure underlying it. Linguistic reference is symbolic reference, which is hierarchically constructed from indexical reference, which in turn arises from iconic reference. The relationship between sense and reference cannot be understood without considering this semiotic hierarchy. This chapter begins with a critical review of the map metaphor in the philosophy of language and philosophy of mind, and argues that it presumes a dyadic model of language, where linguistic symbols map to reality according to some (mathematical or logical) functions. Peirce's classifications of signs is introduced in [Section 5.2](#). The hierarchy of reference is discussed in [Section 5.3](#). [Section 5.4](#) argues how symbolic reference, which has no direct correlations with the world, ungrounds from indexical relationships. [Section 5.5](#) explains why symbols have to be embedded in a certain context to represent. By combining the indexical relationships between symbols and between symbols and objects, symbolic representation regrounds in the world. Finally, the problems of sense and reference are explored within this semiotic framework.

5.2 The map metaphor

A map is a designed representational model of a target area for a specific purpose. The symbols and spatial relationships between them on a map represent corresponding locations and relative spatial distances. The relations and distances between sites or objects are mathematically transformed according to certain projection rules or functions. Different maps are designed for different purposes, and correspondingly, the mapping or projection rules from the area to the map differ. For example, a tourist map is specifically designed to help travelers navigate or explore an area, such as a city, region or tourist attraction. It highlights locations, landmarks and services relevant to tourists while simplifying or omitting irrelevant locations like residential communities. A metro map, on the other hand, represents the routes, lines and stations of a city's metro system. Unlike geographical maps, metro maps are primarily designed for clarity and ease of navigation, often sacrificing geographical accuracy for simplicity and usability.

For philosophers such as early Wittgenstein, logical empiricists, Quine and teleosemanticists, language is a map of the world. Their disagreement lies in what kind of relationships (mapping functions or projection rules) exist between language and the world. This map metaphor of language presupposes a code-like notion. According to this view, language is a sequence of symbols encoding reality according to specific coding rules. Therefore, the tasks for the philosophy of language are to articulate what these rules are (semantics), how they are fixed and how they are used in specific contexts by individuals (pragmatics). I will now take Millikan's conception of the semantic mapping function as an exemplar to illustrate this map metaphor. The reason I choose her theory is that it inherits

Wittgenstein's picture theory of language and reconstructs it in a naturalistic manner. However, I will argue that it is inadequate for dealing with the complexities of reference.

Millikan provides a biological model of language in her biosemantics, a variation of teleosemantics. According to this theory, language is also a kind of biological trait. It has biological functions and is shaped by evolution through natural selection or reinforced learning. Language, or intentional sign systems, has many functions like other biological traits. What is special about language is that some of its functions are successfully performed through the communication of information of interest. This communicative function includes a part called the semantic-mapping function, as defined by Millikan (1984; 2005; 2017). For a language form to successfully perform this part of the function, there must be conditions in the world to support it. Some variations in the language form systematically vary with these conditions (Millikan 2005, 53). Millikan uses the term "mapping function" in a mathematical sense. In mathematics, if there is a mapping function between two sets, each element of one set corresponds to exactly one element of the other set according to some rules that can be mathematically or logically characterized. In the case of the semantic-mapping function, the two sets are language and the world. The elements of the language set are the varieties of language forms, while the elements of the world set are properties, objects, events, structures, processes, etc. The elements of the language set map to those of the world set according to a semantic-mapping function. This function is conventionally stabilized in the speaker-hearer cooperative systems of a language community, with stabilization occurring in relation to its contribution to a conventional function. Furthermore, a semantic function may contribute to different stabilizing functions. Let's consider the example of animal signals: The bee waggle dance.

It is well known that a honeybee communicates the location of a nectar through a figure-eight-like dance. When a bee finds nectar more than 40 meters away from its hive, it returns to the hive and performs the waggle dance. The angle α between the direction the bee moves and the sun indicates the direction of the nectar, and the duration of the wagging, t , indicates the distance from the hive. The location of the nectar can be determined using the direction and distance. As a type of intentional sign system, the semantic-mapping function maps the location of the nectar to the dance. Accordingly, the two sets are the sets of waggle dance and nectar locations. The variables α and t constitute the set of waggle dance, while the locations of the nectar are the elements of the other set. The mapping function in the mathematical sense is the vector relation constituted by α and t , (α, t) . Variables α and t vary systematically with nectar locations. For the bee to communicate the correct and accurate information about nectar location to other bees, when a bee performs a waggle dance (α_m, t_n) , the nectar must be situated in the specific location specified by the dance. The nectar location, in philosophical terms, is the truth condition or satisfaction condition of the waggle dance. The conventional function of the waggle dance is to communicate the information about the nectar's location, and thus, the semantic-mapping function is stabilized with respect to the stabilization of this conventional function.

The linguistic semantic-mapping function is similar to that of a waggle dance, but with a far more complex mapping function and variations. For example, “Immanuel Kant was a German philosopher” represents the state of the world in which Immanuel Kant was a German philosopher. In this simple descriptive sentence, we can identify three variables: the proper name “Immanuel Kant,” the adjective “German,” and the predicate “philosopher.” Each variable corresponds to some aspect of reality. However, this simple sentence can be misleading, as it implies a one-to-one semantic mapping between the components of the sentence and the aspects of the world. In reality, the semantic-mapping functions are much more complex. The variables themselves do not correspond to their objects independently unless each is situated in a sentence. Furthermore, aside from descriptive linguistic representations, whose content is determined by what they need to correspond to, there are directive representations, whose content requires the hearer to take action in order to satisfy it. Speech acts, such as “Attention, please!” are examples of directive representations. For descriptive representations, “it might be the case that the producer, who is primarily responsible for making the sign, corresponds to the world.” These are designed to stand for existing world affairs. For directive representations, “the consumer may be primarily responsible for making the world correspond to the sign.” These are designed to prompt hearers to produce the represented world affairs (Millikan 2004, 80). According to this account, a task of the philosophy of language is to locate the conventional functions of different linguistic representations and their corresponding semantic-mapping functions.

This account of the semantic-mapping function aligns with our intuitions about animal signals and even subpersonal signals. It also works well for explaining indexical sentences and definite descriptions. However, when it comes to sentences that do not have direct referents, such as indefinite descriptions, general descriptions, abstract descriptions and fiction, the theory becomes less clear. Nonetheless, Millikan (1984) has attempted to address these issues. She argues that for an indefinite description like “A philosopher wrote a great philosophical book titled *Critique of Pure Reason*,” its function is not to refer to a specific individual but “only to map onto at least one individual in the context of the whole sentence” (Millikan 1984, 177). That is, indefinite descriptions are supposed to map onto dispositions or kinds, rather than to any specific individual. Similarly, general or abstract definitions like “Animals with kidneys also have hearts” and “The sum of the angles in a triangle is 180 degrees” are also supposed to map onto dispositions or kinds in nature. Unlike Meinong, who develops a metaphysics defining fictional entities as real, Millikan (1984, 201–202) accepts a pretend theory, which asserts that when people use such fictional sentences, they pretend that the represented world affairs exist, even though they do not.

Millikan’s semantic-mapping function theory has the advantage of providing a unified account that covers all intentional signs, from subpersonal signals to animal signals and even to language. However, it is difficult to say the theory is entirely satisfying. First, when she argues that indefinite descriptions refer to dispositions, properties, or kinds, she presupposes a certain metaphysical stance (Millikan 1984; 2017). Dispositions, properties and kinds are universals whose existence has been

debated since ancient Greece. As a result, the conception of the semantic-mapping function carries a metaphysical burden. Second, understanding fictional representation as pretending presupposes that people have the competence to interpret counterfactuals. When someone makes a counterfactual claim, it means they are describing something that does not exist but pretending that it does. This requires higher-order cognitive competence. To understand fictional linguistic representations, the relationship between counterfactual interpretation and language should be explored, but Millikan's theory only assumes it.

Third, Millikan's biosemantics presupposes the continuity between different types of signs, yet it hardly explains the differences between them. For example, in some of her works, she acknowledges the difference between animal signals and language (Millikan 1989a; 2005, 64), but she does not explain where these differences come from. Consider the difference in the story of *The Boy Who Cried Wolf*. With the repeated cry of "Wolf! Wolf!", the indexical power of the sentence is lost. That is, the villagers who hear the boy's cry no longer believe that there is a wolf present (Deacon 1997, 82). However, the meaning, or sense in Frege's terms, of the sentence is not lost. This is also the case for Pavlov's dog. After repeated training, the bell sound becomes a sign for food, acquiring an indicating power to food, but if the trainer stops pairing the sound with food, the bell quickly loses its referential power. The dog no longer interprets the ringing as a sign of food. As a naturalistic theory of meaning, it should not only explain the continuities but also the discontinuities.

Furthermore, although Millikan (2004; 2005; 2017) argues that there is no principled distinction between semantics and pragmatics, there is a tension between the public meaning or conventional functions of language and their individual usage in practice. This distinction arises from the fact that individuals always have their own way of using and understanding conventional language forms. Thus, how an individual employs an intentional sign to identify a world affair in a specific context becomes a crucial problem. Millikan develops a theory of identity and introduces new concepts such as *unicept* and *unitracker* to understand this issue. However, I would argue that the tension reflects the same problem faced by the etiological conception of biological functions analyzed in Chapter 4: The conventional meaning of language forms is a disposition determined by the past. First, disposition is a debatable concept; second, a language form has conventional meaning about the present, not the past. As I have argued in Chapter 4, what evolutionary history provides are constraints on the possible ways a sign can be used, rather than specific functions.

A deep fallacy underlying these map theories of meaning is what is called the fallacy of simple correspondence (Deacon 2003) or the code fallacy (Deacon 2012b). According to the code fallacy, language is a type of code determined by a one-to-one mapping between conventionally determined sign sets (Deacon 2012b, 395). This notion is a fallacy because it presents an oversimplified picture, leading to the difficulties discussed above. The code fallacy overlooks the complex semiotic infrastructure. Symbolic reference in language hierarchically depends on more basic types of reference. I will argue that, with understanding this semiotic infrastructure beneath

symbolic reference, the tension between the public aspect and individual practical usage will disappear; the distinction between sense and reference, along with the relevant problems, can also be articulated. As a result, the functions of various language forms can be explained.

5.3 Hierarchy of reference

In his semiotics, Charles Sanders Peirce defines signs with the presentative, representative, interpretative and triadic conditions, as I introduced in [Chapter 4](#). Accordingly, signs can be classified according to their presentative, representative and interpretative characters: The signs with respect to its ground, the signs with respect to its object, and the signs with respect to its interpretant. What we are concerned with here is the classifications of signs with respect to their presentative and representative characters ([Liszka 1996](#), 35). Moreover, we should keep in mind that signs are triadic in nature, although our focus here is on the referential relationships between the sign and the object, as I discussed in [Chapter 4](#). “No particular objects are intrinsically icons, indices, or symbols. They are interpreted to be so, depending on what is produced in response” ([Deacon 1997](#), 72). What a sign vehicle provides are some constraints shaped by physical, evolutionary, learning, or human intentions. It is the interpreter who interprets the sign vehicle as a sign. Therefore, when we define a sign as a specific type, what we really mean is that it is designed (intentionally by evolution, learning, or some human purpose) to be (highly likely) interpreted in that way.

To analyze the representative characters of signs, we must distinguish them from the presentative characters. As Liszka puts it:

The presentative character of the sign is a certain feature which the sign has qua object (or existent), which forms the basis of its capacity as a sign to represent its object; that is, it serves as the ground of the sign. The presentative character of the sign, then, forms the basis of its representative character.
([Liszka 1996](#), 35)

According to the presentative characters, Peirce classifies signs into qualisigns, sinnsigns and legisigns. For a sign to be a qualisign, its qualitative character serves as the basis for its representative capacity. For example, the redness of a fire, the coolness of ice, and the heaviness of a lead block are qualisigns when these characters serve as the basis for icons. For a sign to be a sinnsign, its existential qualities or its instantiation serve as the basis for its representative capacity. The presentative characters of sinnsigns are particulars in the Aristotelian sense, presenting at a specific time and place. For instance, the instantiation of smoke at a certain location and time is a sinnsign. For a sign to be a legisign, the character serving as the basis for its representative capacity is conventional. That is, the presentative character of a legisign has nothing to do with the intrinsic properties of its sign vehicle. For example, what makes the word “dog” a legisign has nothing to do with the physical, structural or spatial-temporal properties that realize the word.

Now, let's turn to the topology of signs according to their representative characters. According to the manner in which a sign correlates with its object and thus represents that object, Peirce classified signs into icons, indices and symbols. If a sign is an icon of its object, the presentative character of the sign is similar to some character of that object. A portrait is an icon of a person because of the shared appearance features. A tourist map is an icon of a tourist attraction because the spatial relationship characters presented on the map are similar to those of the place. If a sign is an index of its object, then its presentative character is contiguous with that of the object. When an object has a spatial, temporal or causal relationship with another object, they are contiguous, and thus, one can be interpreted as an index of the other with respect to the contiguous characters. For example, smoke is an index of fire because it is caused by and coexists with fire in a contiguous space. A windvane is an index of the wind because its flying direction is parallel to the wind, and its shape, blown by the wind, is spatially isomorphic to the wind. The bell sound is an index of food for Pavlov's dog because it always coexists with food spatially and temporally. When a sign is a symbol of an object, its correlation with the object is conventional. A correlation between two objects being conventional means that it is arbitrary, having nothing to do with their intrinsic properties. Linguistic symbols are a typical type of symbols. But symbols are not limited to language; various religious symbols and organizational logos are symbols, too. For instance, the cross is a symbol of Christianity and the Red Cross is a symbol of medical aid.

Combining the classifications of signs in virtue of both their presentative and representative characters, we have six types of signs: Iconic qualisign, iconic sinn-sign, iconic legisign, indexical sinn-sign, indexical legisign and symbolic legisign. A qualisign cannot be an index because the representative character of an index involves spatial, temporal or causal relationships between a sign and its object, which are particulars, whereas qualitative characters are not particulars. A qualisign cannot be a symbol because its presentative character is not conventional. A sinn-sign cannot be a symbol because its presentative character is a particular instantiation at a specific time and space, which is not conventional and thus cannot serve as the basis for a conventional relationship.

Traditions in the philosophy of language usually overlook the distinction between presentative and representative characters, thereby confusing different types of signs. As a result, various referential relationships are erroneously identified as a single type. Animal signals, like the bee waggle dance, are indexical sinn-signs. The presentative characters of a waggle dance, serving as the ground of the sign, are the duration of the wagging and the bee's head direction. A waggle dance is a world affair happening at a specific time and place. The presentative characters of the waggle dance are contiguous with those of the location of nectar: The duration of the wagging and the direction of the bee's body vary with the distance and direction of the represented nectar. A map is an iconic legisign. It is a legisign because the presentative characters of the map, serving as the ground of the sign, are conventional. Geometric figures and elements are conventional; what matters to their readers has nothing to do with their specific physical realization of them. A map is an icon because of the similar structural

relationships it shares with the represented territory. Since the grounds of different types of reference (the presentative characters) are different, we should not expect a sole type of referential relationship to explain all types of reference. Therefore, we should not take animal signals and language as the same type of signs. Language is not strongly continuous with animal signals. That is, they do not share a common abstract pattern. The functional characters that are fundamental to animals are not fundamental to language.¹ Nevertheless, this does not imply that different types of reference are independent of each other. It is hard to understand how the function of symbolic reference is performed in practice if this were the case. Rather, they are asymmetrically dependent. Symbolic interpretation depends on indexical interpretation, which, in turn, is constituted by iconic interpretation. Let us use the example of the bee waggle dance to illustrate the index's hierarchical dependence upon the icon.

The first type of iconic characters beneath the waggle dance is the isomorphic relationship between the presentative characters of the waggle dance and the location of nectar: The angle varies isomorphically with the relative spatial relationship between the wagging bee, the nectar and the sun; the duration varies isomorphically with the distance from the hive to the nectar. Recall that it is other bees around the wagging bee who interpret the waggle dance as an index of nectar. Therefore, these other bees must have the interpretive capacity to recognize the resemblance between the two sets of presentative characters. The second type of iconic character builds on the resemblance between the presentative characters of the current waggle dance and those of past waggle dances, as well as the present nectar and the past nectars. The third type of iconic character is the persistent correlation between the present characters of past waggle dances and past nectar locations. Only by acquiring at least these three types of iconic characters can a bee grasp the capacity for indexical reference.

Symbolic reference also hierarchically depends on indexical reference. An example from archaeology provided by [Deacon \(1997, 72\)](#) clearly illustrates this hierarchy. An archaeologist discovers some clay tablets with markings. Since these markings obviously exhibit some regularities resembling other linguistic symbols, it is easy for the archaeologist to speculate that they may be linguistic symbols. However, since the archaeologist does not know which language these markings belong to, she or he cannot interpret them symbolically. Nevertheless, the archaeologist can interpret the markings as iconic legisigns. Later, the archaeologist realized that the same type of markings were found along with trade goods that had been transmitted from one place to another. Some imprints repeatedly appear with the same goods. This provides the archaeologist with the iconic infrastructure in a specific context to interpret the markings as indexical legisigns. These markings may indexically refer to the goods with which they are associated. With more regularities found between these indexical legisigns, the archaeologist may eventually decode the markings and acquire a symbolic interpretation.

These two examples show that the hierarchy of reference reflects the hierarchical dependence of different levels of interpretation. As Deacon states,

It's not just the case that we are able to interpret the same sign in different ways, but more important, these different interpretations can be arranged in a sort of ascending order that reflect a prior competence to identify higher-level

associative relationships. In other words, reference itself is hierarchic in structure; more complex forms of reference are built up from simpler forms. But there is more to this than just increasing complexity. This hierarchical structure is a clue to the relationships between these different modes of reference. Though I may fail to grasp the symbolic reference of a sign, I might still be able to interpret it as an index (i.e., as correlated with something else), and if I also fail to recognize any indexical correspondences, I may still be able to interpret it as an icon (i.e., recognize its resemblance to something else). Breakdown of referential competence leads to an ordered descent from symbolic to indexical to iconic, not just from complex icons, indices, or symbols to simpler counterparts. Conversely, increasing the sophistication of interpretive competence reverses the order of this breakdown of reference.

(Deacon 1997, 73–74)

We have argued that indexical interpretation asymmetrically depends on the iconic. But how does symbolic interpretation depend on the indexical? Unlike icons and indices, whose presentative characters of the sign vehicle directly correlate with those of the referent, symbols have no direct correlation with their referents. So, how does the cognitive competence of symbolic interpretation arise from the competence of indexical interpretation? This is the problem of symbol ungrounding: how does the symbol free itself from a direct correlation with the world or reality? Moreover, humans do use language to talk about objects, properties, events and more. In other words, symbolic representation has indexical competence. While symbolic reference does not directly correlate with the world, it can regain the capacity for direct reference in some way. For example, the sentence “The 47th president of the United States has claimed that Canada should join the USA to become the 51st state” refers to a specific world event. This is crucial for understanding how language functions in specific contexts.

Furthermore, iconic and indexical references resurface within linguistic symbolic systems. That is, one language form may have iconic or indexical correlations with other forms with respect to their syntax and semantics. The hierarchy of reference is recreated in linguistic and symbolic systems. Solving the symbol ungrounding and regrounding problems can provide us with a clearer picture of how language functions, helping to resolve the difficulties faced by the mapping metaphor of language.

5.4 Symbolic reference is different

In his classical paper “The Symbol Grounding Problem,” Stevan Harnad (1990, 335) addresses a challenge faced by artificial intelligence at the time:

How can the semantic interpretation of a formal symbol system be made intrinsic to the system, rather than just parasitic on the meanings in our heads? How can the meanings of the meaningless symbol tokens, manipulated solely on the basis of their (arbitrary) shapes, be grounded in anything but other meaningless symbols?

In short, he explores how symbols can connect to the world. The concept of a symbol Harnad uses refers to arbitrary tokens, which is different from the concept we use in this chapter. It corresponds to a legisign in Peirce's semiotics. Harnad argues that there is a fundamental difference between the symbol tokens used in artificial intelligence and those in language. The former are manipulated purely based on their formal or syntactic properties, whereas the latter are inherently connected to the world. Therefore, the real problem of artificial intelligence is how to make legisigns intrinsically meaningful.

To understand how symbolic reference is possible, we must first understand what distinguishes it from indexical reference and then explain how its distinctiveness is constructed from the indexical infrastructure. Let's briefly revisit the distinctive features of symbolic reference mentioned earlier. First, symbolic reference is conventional. The conventionality of symbolic reference has two aspects: The conventionality of the sign vehicle and that of the referential correlation. By defining legisign and symbol according to their presentative and representative characters, we distinguish them as two types of conventionality. The full understanding of conventionality will be discussed in [Chapter 6](#). Second, symbolic reference has no direct referential correlation with its possible referents and enjoys a certain degree of autonomy. The story "The Boy Who Cried Wolf" demonstrates how the indexical function of the sentence "Wolf! Wolf!" is lost because it no longer correlates with a real wolf in those situations, while the symbolic reference remains ([Deacon 1997](#), 82). Third, symbols are capable of representing something general, abstract or fictional. As introduced in the first section, the relationship between sense and reference and how to handle representations with no real referents have been persistent problems in the philosophy of language.

Fourth, a symbol must be embodied in a specific context to have a specific referent, even for proper names. For example, the word "dog" has a dictionary or conventional meaning, such as the definition from the Merriam-Webster Dictionary: "a carnivorous mammal (*Canis familiaris*) closely related to the gray wolf that has long been domesticated as a pet." However, it refers to no specific dog but rather to an indefinite description. Suppose a person is walking down the street and sees a dog. If she says, "Dog!" then "Dog!" refers to the dog she sees. The word gains its indexical function by being embedded in this specific situation. For the hearer, what constitutes the sign of the dog is not just the word itself but the entire physical context, including the spatial-temporal location of the person who speaks the word. Thus, the hearer interprets "Dog!" as an indicator of the dog by interpreting the whole situation in which the word is spoken. For example, in the sentence "The first living creature to orbit Earth was a dog," the word "dog" refers to Laika, the dog sent into orbit aboard Soviet spacecraft Sputnik 2 in 1957. Here, the word "dog" fixes its referent by embedding itself in the sentence.

Furthermore, some may argue that, unlike other linguistic symbols, proper names have a direct indexical function. For instance, "George Washington" unambiguously refers to the first president of the United States. Nevertheless, as Saul Kripke's causal theory of reference (1980) posits, the referential function of proper names is constructed through a causal-historical chain. A proper name refers to a

referent due to a series of causal interactions that began when the referent was first identified and named. The referent of a proper name remains as long as the causal chain persists. However, the causal theory cannot fully explain the second feature of symbolic reference.

Fifth, unlike indexical references, which are independent of one another, language is organized according to a universal grammar that enables its systematicity and generativity. For example, vervet monkeys have three distinct signals for three types of predators. These signals are unrelated to one another. No need to say they are meta-signals about the signals. Language, on the other hand, has different components that combine systematically according to certain rules. Letters compose words, which compose phrases, which compose sentences, which compose paragraphs and so on. This generates complex linguistic structures. Language can generate nearly infinite sentences due to the various possibilities for combining components. This generativity of language is considered an essential characteristic by modern linguists, particularly [Chomsky \(1968; 1980\)](#).

These distinctive features of symbolic reference present some of the most persistent problems in the philosophy of language and linguistics. However, these features are superficial. Focusing solely on them overlooks the rich semiotic infrastructure underlying symbolic reference. Following Deacon's interpretation of Peircean semiotics, I argue that these features are products of a rich semiotic infrastructure and are deeply interconnected. As introduced in the previous section, more basic referential relationships—namely, icon and index—form the foundation from which symbolic reference is built. Symbolic reference hierarchically depends on the index, which in turn depends on the icon. Accordingly, Deacon and his colleague Joanna Raczaszek-Leonardi suggest that we should invert the symbol grounding problem:

...the problem is actually to explain how iconic and indexical forms of communication—which are intrinsically “grounded” due to the sign vehicles sharing features with their referents—can be used to develop communication using ungrounded sign vehicles (aka words/symbols).

([Deacon 2018](#), 84)

Just as indexical reference is constructed from the iconic infrastructure; symbolic reference is similarly constructed from the indexical. The problem, then, is how symbolic reference depends on indexical reference, becomes detached from direct reference, and eventually becomes ungrounded. The distinctive features of symbolic reference offer clues to solving this problem.

5.5 Symbol ungrounding

In the Peircean semiotic framework, as I have discussed above, it is not the intrinsic characteristics that define the type of sign but the interpretive process. So, what kind of interpretive process is involved in interpreting a legisign as a symbol? Unlike indexical reference, which is independent of each other and constituted by

singular properties, processes or events (as I've discussed earlier), a legisign embeds in a complex system of other legisigns upon which its representational power depends. A symbolic representation, such as a sentence, may lose its direct referential competence, as illustrated by *The Boy Who Cried Wolf*. However, it does not lose its potential representational power. The reason symbolic reference can separate its competence of referring from its actual performance is that "it maintains its fluid and indirect referential power by virtue of its position within a structured set of indexical relationships among symbol tokens" (Deacon 2003, 122). This means that its representational power is distributed across the stable relationships between words (Deacon 1997, 83).

Although a symbolic representation has meaning or "sense" in Frege's terms, on its own, its representational power depends on its connection to other symbols. This combination of symbols forms an indexical system in which one symbol indicates another. For example, in the sentence "A dog has four legs," the word "A" is an indicator of the word "dog" because of the spatial continuity between them. The sentence is an indefinite description and thus lacks a definite referent, but it still has meaning. This implies that a symbolic system is actually a distributed indexical network, where a legisign has indexical correlations with other legisigns. These indexical correlations are manifested as structural relationships, such as grammar or syntax, between legisigns. As a result, symbolic reference has a degree of autonomy. Unlike icons and indices, the representational power of symbols does not rely on direct, particular correlations between the sign vehicle and its object. The distributed indexical network and the resulting autonomy suggest that the interpretive strategy has shifted from particular indexical relationships to abstract relationships between legisigns. When interpreting a legisign as a symbol, one no longer concerns oneself with the direct correlation between it and its object, but rather with the relationships between it and other legisigns.

This transformation does not involve learning something entirely new or applying past experience to new situations. This is what indexical reference does—finding resemblances between past and present. Instead, it is a transformation of interpretive strategy, what Deacon (1997, 92) calls an "unlearning insight." Symbolic representational competence, then, is not learned but discovered. This can be illustrated through the example of learning a new language.

When learning a second language, we often start by memorizing simple, frequently used expressions like "Hello," "Thank you," and "How are you?" We must then remember many words and phrases. However, we do not directly learn the second language; we use our native language as a medium. We interpret the meaning of the words, phrases and sentences by translating them into the equivalent terms in our native language.

To grasp the second language, we work hard to remember a lot of words, phrases and grammar through their translation. This learning process is not symbolic interpretation but rather indexical interpretation in the early stages of language learning. The list of words and phrases in a textbook presents the indexical correlations between words and phrases in the second language and their counterparts in the native language. A word or phrase in the second language corresponds to some word

or phrase in the native language. In the second stage, we may realize that there is a kind of systematic mapping between these two languages. The structural relationships between legisigns in one language provide a mapping function for the native language. However, aside from the considerable amount of memorization required, one difficulty learners face is that there is often no precise or linear correspondence between words and phrases in two languages. Instead, the mapping is complex and indirect. For example, there is no direct translation of “bread” into Chinese, as bread did not exist in China historically. Similarly, “sky” and “heaven” do not precisely correspond to “天” (tiān) in Chinese.

Once we grasp the second language, we shift our interpretive strategy. We stop translating word-for-word and instead recognize the systematic relationships between words and phrases within the second language. We no longer rely on the indexical relationship between the second language and our native language. Rather, the meaning of a legisign (word or phrase) becomes a function of its relationship to other legisigns in the system (Deacon 1997, 86). At this point, the second language is ungrounded from our native language in our interpretation. When we encounter a new word, we can quickly incorporate it into the system of the second language rather than grasp it through its indexical relationship to our native language.

In learning a second language, we are already equipped with symbolic interpretive competence. But this example seems to postpone the question: How do we acquire our native language in the first place? A notable phenomenon in language learning is that infants acquire language rapidly, usually around two years old. Before this age, there was no systematic learning like we experience when learning a second language, but instead, a variety of interactions between the infant and caregiver(s).

Given this phenomenon, Chomsky (1968; 1980) develops the poverty of stimulus argument. According to this argument, there is insufficient linguistic input to explain how infants acquire language. Linguistic interactions between an infant and caregiver are often fragmentary, unstructured and grammatically incorrect. There are a few corrections for linguistic mistakes. Nonetheless, infants usually acquire complex grammar and near-infinite generative ability, speaking sentences they have never heard before effortlessly. Since language cannot be learned from such input, it must be innate—part of a universal grammar (Chomsky 1968; 1980; Pinker 1994).

However, Deacon (1997) argues that there are no significant anatomical or physiological differences between the human brain and that of other animals. We cannot find anything resembling a linguistic faculty. He and Raczaszek-Leonardi (forthcoming) argue that the Universal Grammar approach reverses cause and effect. The universal structural regularities in language result from symbolic interpretation, not the reverse. Now, let’s explore how an infant learns language. In fact, they learn it in a way similar to how we learn a second language.

Unlike traditional developmental linguists, who claim there is insufficient linguistic input during infancy, infants actually experience rich semiotic interactions with their caregivers. Simple expressions like “Mum!,” “Dad!,” and “Milk!” along with various sounds, facial expressions and eye contact, serve indexical functions.

These expressions have spatial-temporal or causal continuities with the represented properties, processes or events. What distinguishes these indexical correlations from those in translation is that they are between legisigns produced by the caregiver and objects. These indexical legisigns are no different from animal signals, such as vervet monkey warning signals or a bee's waggle dance. We should bear in mind that, what I mean here should be understood in Peircean triadic framework of semiotics. That is, infants are capable of interpreting these legisigns as indexes just like animals.

However, infants are distinct from animals in their cognitive abilities. As their brains and related cognitive faculties develop, and as they acquire numerous indexical relationships, infants begin to recognize the relationships between legisigns apart from their direct correlations with objects. With this recognition of legisign-to-legisign relationships, there is a shift in mnemonic strategy, from relying on indexical relationships between legisigns and objects to recognizing relationships between legisigns that indirectly fix the referents. As a result, a symbol's representational power depends on its position within an interlinked symbol network. Moreover, its representational power no longer depends on its physical correlation with the object, as it becomes distributed over the system and gains autonomy from the reality it represents.

This transformation reduces the burden on memory and makes logical or categorical generalization possible (Deacon 1997, 88). When learning a new symbol or word, the referential relationship is no longer a function of continuity between the legisign and object. Instead, it becomes a function of the relationships this new word shares with other words in the system, which offers a limited set of ways to integrate new items (Deacon 1997, 88).

As Deacon (1997, 99–100) says,

In summary, then, symbols cannot be understood as an unstructured collection of tokens that map to a collection of referents because symbols don't just represent things in the world, they also represent each other. Because symbols do not directly refer to things in the world, but indirectly refer to them by virtue of referring to other symbols, they are implicitly combinatorial entities whose referential powers are derived by virtue of occupying determinate positions in an organized system of other symbols. Both their initial acquisition and their later use requires a combinatorial analysis. The structure of the whole system has a definite semantic topology that determines the ways symbols modify each other's referential functions in different combinations. Because of this systematic relational basis of symbolic reference, no collection of signs can function symbolically unless the entire collection conforms to certain overall principles of organization. Symbolic reference emerges from a ground of nonsymbolic referential processes only because the indexical relationships between symbols are organized so as to form a logically closed group of mappings from symbol to symbol. This determinate character allows the higher-order system of associations to supplant the individual (indexical)

referential support previously invested in each component symbol. This system of relationships between symbols determines a definite and distinctive topology that all operations involving those symbols must respect in order to retain referential power.

Moreover, indefinite, general and abstract descriptions are the results of symbolic interpretation. Due to the distributed indexical power and autonomy of symbols, the meaning—or Fregean sense—of these representations is ungrounded in physical reality, that is, they do not necessarily have physical referents. There is no need to worry about the metaphysics of their referents, as they may not have corresponding physical referents despite what the content of these representations describes. For an indefinite description, there could indeed be corresponding referents as described, but the description does not directly represent them. For example, “A dog was named Bark” is an indefinite description. It is entirely possible that there exists a dog named Bark, but the description itself has no direct correlation with that dog. Instead, it acquires its content through the semiotic system beneath it.

The same holds true for general and abstract descriptions, though they differ slightly. Unlike indefinite descriptions that may refer to a specific object, general and abstract descriptions serve as logical or categorical generalizations. These generalizations reflect the transformation of an interpreter’s mnemonic strategy, which shifts the burden of remembering numerous singular indexical correlations to the symbolic relationships themselves. For example, “A dog has four legs and barks” is a general and abstract description. It categorically generalizes certain properties of dogs without requiring one to remember the specific attributes of all particular dogs.

Consequently, this framework allows us to semiotically understand Frege’s distinction between sense and reference, shedding light on how indefinite, general and abstract descriptions challenge traditional theories of reference and meaning. It also clarifies the continuities and discontinuities between animal signals and language, or symbolic representation, by highlighting the asymmetrical dependencies between icon, index and symbol. The next question then becomes: How does a symbol determine its definite referent, given that its representational power is distributed across the symbol system and lacks a direct correlation with its referent? This is the task of the following section.

5.6 Symbol regrounding

As we can see, only icons and indexes have direct referential relationships, while symbols enjoy some autonomy independent of the referent. So, how can a symbol acquire its capacity to refer to a specific referent? What is the relationship between sense and reference in the Fregean sense? As I argued in the previous sections, the autonomy of symbolic reference depends on the indexical relationships between legisigns. This autonomy makes indefinite descriptions, fiction, abstract concepts and general descriptions possible. But then, how does a definite description determine its represented object?

To acquire its representational power, a symbolic representation requires a second kind of indexical correlation that indicates the objects of reference. For example, when someone says, “That dog has four legs,” in a specific context with a definite dog, the sentence has a representational power by referring to the state of the dog. When the person says this, there must be a way for her to indicate the dog directly, such as by using a pointing gesture or a specific angle of view to fix the referent of “that,” which is a demonstrative pronoun. With this indexical correlation, the sentence gains its representational power. Therefore, there are at least two kinds of indexical correlations underlying symbolic reference. In a specific context, a legisign points to other legisigns, on the one hand, and points to the object of reference, on the other hand (Deacon 2003, 123). Through the interdependence of these two types of indexical relationships, a symbol acquires its representational power. This interdependence is realized through the transitive ability of the indexical relationships.

The transitivity of the index’s representational power means that an indicator of an indicator of an object can also be interpreted as an indicator of the object itself. For example, the sign vehicle of an indexical sinnsign is a natural event that is spatially or causally contiguous with its represented object. The events before and after the sign vehicle can also be interpreted as indicators of the object. For instance, the fire and smoke of a fire beacon serve as an indicator of an invasion. The actions of lighting the beacon fire and the smell of the fire can also be interpreted as indicators of the invasion. Unlike sinnsigns, the sign vehicle of symbols is legisigns. Similar to sinnsigns, when a legisign is interpreted as an indicator, its representational power is also transitive. Its transitivity is realized by its position in a symbol system and the spatial relationships it has with other legisigns. For example, in the sentence “A dog has spots. It likes to chase cats. A cat comes from Asia. It is a leopard cat,” the first “it” refers to the dog, while the second “it” refers to the cat. We judge this based on their spatial relationships: The first “it” is adjacent to the word “dog,” while the second “it” is adjacent to “cat.”

To properly perform their functions, various words—such as articles, nouns, pronouns, verbs, adjectives and adverbs—rely on these spatial relationships. It is intuitive to interpret adjacent relationships between words as indexical. For example, the adjacent relationships between words in the sentence “A dog has spots” form a combination where each word indicates the others. Some word types, like prepositions and conjunctions, also perform indexical functions, similar to direction arrows on roads. For example, in the sentence “A cat and a dog come from Asia,” the conjunction “and” links the two words that are spatially continuous with it, while the preposition “from” spatially links the subject “a cat” and “a dog” to Asia. The prepositions “and” and “from” in this sentence function like direction arrows pointing to words contiguous to them. Based on this, we may predict that, since words are adjacent to each other in a sentence, implying spatial relationships, prepositions may not always be necessary. They support verbs and adjectives in realizing their functions. In fact, there are no prepositions in ancient Chinese. Verbs perform all the functions that are realized by combinations of verbs, adjectives and prepositions.

Linking verbs, including “being,” also perform similar indexical functions, correlating the subject and predicate in a sentence. For example, “turn” in the sentence “The sky turns dark before the storm” is a linking verb that connects the subject “the sky” to the predicate “dark.” “Is” in the sentence “A dog is a mammal” links the subject “a dog” to the predicate “a mammal.” These linking verbs function within a distributed legisign network and do not have simple corresponding physical referents. The metaphysics of “being” may be one of the most difficult problems in the history of philosophy. Within our framework, it may be a result of the coding fallacy. Existing linking verbs have no simple corresponding referents; their referents are the parts of the sentence they link, not something outside of the sentence. The sentence as a whole represents something external if there is further index linking it to the world. For example, the pronoun “that” in the sentence “That dog is barking” links the sentence to a specific dog in the context in which the sentence is expressed. This will be discussed in more detail later.

With the systematic regularity of these indexical relationships, a language can be interpreted as a multidimensional vector coordinate system (Deacon 2003, 123). We can understand this system analogously to mathematical functions. The regularity of these indexical relationships is akin to mathematical mapping relationships, with the types of words acting as variables. Just like a mathematical vector coordinate system, where the value of a variable can be fixed, it forms a space of potential reference in which the reference or function of every legisign is defined and determined. As a result, on the one hand, a language has its own autonomy of meaning independent of reality; on the other hand, for a sentence to represent something in the world, it must correlate with reality. This is where the second type of indexical correlation comes in, realized by words like proper names, articles and pronouns.

The referents of proper names are definite due to their causal-historical chain, as Kripke has argued. This provides a point where a sentence with a proper name represents an object in the world. For example, in the sentence “George Washington was the 1st president of the United States,” “George Washington” is a proper name referring to the person George Washington. The referent of the proper name is fixed by its causal-historical chain—perhaps George Washington’s parents named him “George Washington,” and since then, people have called him by that name, and he acted under that name, etc. When a proper name is used in a sentence, this causal-historical chain is transmitted to the indexical relationships between words. The meaning of other words in the sentence indexically refers to what the proper name represents. For example, in the sentence “George Washington was the 1st president of the United States,” the proper name “George Washington” is linked to the description “the 1st president of the United States” through the linking verb “was.” Thus, its causal-historical chain is also transmitted to the description. This analysis shows that the indexical power can be transitive between sinnsigns and legisigns. In fact, the transitivity of indexical relationships between sinnsigns and legisigns serves as a junction between the two kinds of indexical correlations that contribute to determining the referents of symbolic representations. This is clearly demonstrated by pronouns and articles.

In a sentence, a pronoun replaces a noun or object in context to avoid repetition and improve sentence flow. For example, “it” in the sentence “A dog has spots. It likes to chase cats” refers to “a dog.” This is an indefinite description. In a situation where two people encounter a dog, one speaks to the other with a gesture pointing to the dog, saying, “That dog has spots.” The word “that” is a demonstrative pronoun, spatially contiguous with the speaking person and her gesture, thus fixing the referent of the sentence: The dog in this specific situation. Personal pronouns are a bit more complex. “I” always represents the person saying the sentence, and “you” always refers to the person the speaker is addressing. As for third-person pronouns, they can refer to a definite or indefinite person, depending on the context. I will not provide a comprehensive analysis of the indexical functions performed by all types of pronouns, but readers interested in this can explore it further and test this semiotic theory of language.

To perform their proper function, interjections and single-word sentences always require spatial-temporal or causal continuity with sinnsigns to fix their reference. Their function depends on the situation in which they are embedded. For example, imagine a situation where two people encounter a dog. One speaks to the other with a gesture pointing to the dog, saying, “Look! That dog has spots.” The single-word sentence “Look!” is embedded in this situation. It indexically (spatial-temporally) correlates with the position of the speaking person and the direction of her gesture. The hearer understands that “Look!” refers to the direction the speaker is pointing to. A sentence with an interjection, such as “Wow! How beautiful it is!” must also be spoken in a situation. The continuity between the sentence and the context also determines what the pronoun “it” stands for.

Now, let’s turn to the more challenging aspects of language, such as fiction and metaphor. In fictional descriptions, non-existent things are described. How do we understand them? Within the semiotic theory we have developed, fictional descriptions have no indexical correlation with the real world. In other words, compared to sentences that have determinate real objects, fictional descriptions lack a connection with sinnsigns in reality and thus are independent of it. For example, “Monkey King Sun Wukong causes Havoc in Heaven” represents a fictional figure, Sun Wukong, from Chinese mythology and a fictional event: A havoc in Heaven caused by Sun Wukong. This description has no connection to reality because neither Sun Wukong nor the event exists. The question arises: What do they represent? How can they represent something? When people use words metaphorically, they often do so in ways that stray far from their original meaning. For example, “The stone is very hard,” “Your heart is hard,” and “This building is a product of USSR hard functionalism.” These three sentences use the same adjective, “hard.” However, what they represent is different in each case, though there are some similarities. In the first sentence, “hard” refers to the physical hardness of the stone. In the second, it suggests that the person lacks empathy, has emotional resilience, or possesses an unforgiving nature. In the third, it refers to a building style focused on function over aesthetics, minimalism and industrial design, with a rejection of ornamentation. While there are similarities in these uses, they are indirect and weak. So, how is metaphor possible?

A distinctive feature of symbolic reference is that, as argued above, a symbol may not represent any specific object in reality but does not lose its potential representational power. The reason is that the representational power of a symbol is distributed through its indexical relationships with other symbols in a symbol-to-symbol network. As a result, when a symbol loses its direct indexical correlation with any object and becomes ungrounded from reality, it does not mean the iconic infrastructure beneath the indexical relationships is also lost. The iconic relationships are still preserved in the interpreters. A symbolic system as a whole shares iconic relationships with the real world. Moreover, symbolic representation makes logical or categorical generalization possible. This type of generalization depends on the similar properties shared by different objects—an iconic relationship! That is, people interpret different objects, events, or processes as the same and thus classify them into categories. For example, different particular tigers are treated as members of a species due to their shared properties. Nouns, verbs, adjectives, adverbs and interjections are categorical generalizations in this sense. This type of iconic relationship further supports the indexical relationships between symbols and between symbols and objects, as analyzed in [Section 5.2](#). Fictional descriptions and metaphors also rely on this type of iconic relationship.

Even though there is no referent in reality for a fictional description, there are similar properties shared by the description and some real objects. For instance, there is no event in reality represented by the fictional description “Monkey King Sun Wukong causes Havoc in Heaven.” However, when someone reads this description, they interpret it based on its resemblance to real objects and events. Though the nouns (such as the proper name “Sun Wukong”) and verbs are categorical generalizations, what they represent is fixed by the indexical relationships between them. Thus, the description represents a definite fictional event. Metaphors work in the same way. For example, while a stone is completely different from a person’s mental state, there can be similarities between the properties of the stone and those of the mind. Saying “A stone is hard” means it is difficult to break or penetrate, while saying “Someone’s heart is hard” implies that their emotions are difficult to change. The similarity between the stone and the person’s mental state lies in the concept of “difficulty to change.” This type of iconic relationship supports the indexical relationships between symbols, and symbols and objects. As a result, “Thus the combinatorial organization of symbolic legisigns comprising a phrase, sentence, or narrative constitute a higher order iconic, indexical, or symbolic referential function”(Deacon 2012b, 400).

Finally, let’s revisit the map metaphor in traditional theories of representation. Although there are indeed some iconic relationships between a symbolic system and the world, traditional notions oversimplify this relationship because they overlook the rich semiotic infrastructure beneath the superficial correspondence. The picture of language provided by this semiotic framework supports a kind of holism. Language as a whole has the power to represent the world, but there are no simple, corresponding relationships between them. Language regrounds in the world through two types of interdependent indexical relationships: Symbols point to the objects of reference while also pointing to other symbols in the system.

5.7 Conclusion

In traditional naturalistic models, representation is often viewed as a kind of semantic mapping from the representation to the represented. However, this mapping conception suffers from what is called the “code fallacy,” which makes it difficult to explain the variety of linguistic phenomena. Peircean semiotics offers an alternative naturalistic approach that better accounts for these complexities. According to Peirce’s triadic model of representation, there is a rich semiotic infrastructure underlying language. Semiotic interpretation operates asymmetrically: Symbolic interpretation depends on indexical interpretation, which in turn depends on iconic interpretation. This framework offers a novel understanding of the continuities and discontinuities between animal communication and human language. With this semiotic infrastructure in place, symbolic reference becomes detached from direct correlations with objects and instead regrounds in the world through two interdependent types of indexical relationships. This structure helps explain phenomena such as indefinite, fictional, abstract descriptions and metaphors.

Note

- 1 This expression is revised from Peter Godfrey-Smith’s life-mind strong continuity thesis, “Strong Continuity: Life and mind have a common pattern or set of basic organizational properties. The functional properties characteristic of mind are an enriched version of functional properties that are fundamental to life in general. Mind is literally *life-like*” (Godfrey-Smith 1996).

6 Convention

6.1 Introduction

We use the word “dog” to refer to the animal dog and “cat” to refer to a cat. It is logically possible, however, to use “cat” to represent a dog and “dog” to represent a cat. This is because there are no natural correlations between the word “dog” and the animal dog, or between the word “cat” and the animal cat. Their relationships are arbitrary. It is convention that makes “dog” represent a dog. This example highlights a distinctive feature of language: Conventionality. It is widely believed that conventionality makes such arbitrary representational relationships possible. Inspired by research in animal communication and signaling, the naturalistic approach to representation and meaning argues that arbitrariness and conventionality are also characteristics of animal signals. For example, vervet monkeys use different alarm calls to signal the presence of different predators. There are no inherent connections between the intrinsic properties of the signals, such as frequency and duration, and the characteristics of the predators. These findings in ethology provide clues for developing a unified account of intentionality, meaning and representation (Millikan 2005; Skyrms 2010a). Therefore, if conventionality can be explained, the emergence of meaning can be explained as well.

However, W.V. Quine (1936) argued that this notion leads to a vicious circle. Conventions rely on agreements, and in order to make an agreement, we must be able to communicate through some form of primary sign system—yet the emergence of sign systems is exactly what we aim to explain. David Lewis, whose PhD was advised by Quine, challenged Quine’s argument in his dissertation and then published as *Convention: A Philosophical Study* (1969). He argued that conventions emerge from social interactions between different agents and formalized the process as signaling games. It was argued that signaling systems in which information is communicated from the sender to the receiver in a simple signaling game are strict Nash equilibria in the game. Then, the problem of the emergence of meaning becomes the problem of how to converge to and maintain strict Nash equilibria in signaling games, a coordination problem in game theory. The solutions provided by Lewis are common knowledge and salience. However, Brian Skyrms (1996; 2004) argues that Lewis’ solution cannot escape Quine’s critique. Instead, Skyrms proposes an evolutionary dynamic approach to the problem. The dynamic analysis

of signaling games shows that signaling systems spontaneously emerge in the interactions between senders and receivers. Common knowledge and salience are unnecessary.¹

It is believed by many philosophers that the framework of Lewis-Skyrms signaling games brings fundamentally new insights to questions concerning the explanation of meaning in general. As Birch summarizes,

First, he moves from classical to evolutionary game theory: the convention-establishing work that is done by conscious rational choice in Lewis's models is done by natural selection or trial-and-error learning in those of Skyrms. Second, Skyrms moves from purely static analyses based on the concept of a coordination equilibrium to dynamic modelling of how signalling systems evolve over time. Third, Skyrms combines his game-theoretic models with an information-theoretic account of what it is for a signal to carry informational content: an account that can be seen as a generalization of that of Dretske (1981).

(Birch 2014, 497)

This reformation endows signaling games with several virtues. First, it does not need agents to be rational in any sense (Huttegger 2014). Signaling emerges spontaneously without belief. The virtue makes signaling games a general theory of signaling interactions. It is not just limited to human signaling but can be applied to explain signaling interactions between and within organisms (Harms 2004; Godfrey-Smith 2014). The second virtue is that the dynamics of various kinds of signaling games can be deeply and extensively studied with the help of computer simulation. It makes the framework of signaling games a powerful theoretical tool. Fruitful works in this respect have been produced in the last decade (Skyrms 2010a; Huttegger et al. 2010; Huttegger and Zollman 2011; Huttegger 2014, to name a few). Third, combined with information-theoretical ideas, the conception of signaling games becomes an insightful way of approaching a general naturalistic theory of meaning or content (Huttegger 2007a; Birch 2014; Shea et al. 2018; Skyrms and Barrett 2019; Godfrey-Smith 2020). As a powerful tool, researchers also apply it to relevant topics like the emergence of language (Huttegger 2007b; Barrett 2009; Franke and Wagner 2014; Steinert-Threlkeld 2016, 2019; Barrett and LaCroix 2020), logic and truth (Barrett 2016).

In the chapter, I will argue that mutual misunderstanding may exist in interactions between epistemic agents. I present a Chinese folk story, *Magical Fight*, as an exemplar of mutual misunderstanding and a formal definition of mutual misunderstanding. The story shows that although the interactions between the two players in the magical fight are successful for both players and the audiences of the fight, there is no effective communication between the players. In the case of mutual misunderstanding, players reach an equilibrium while their interpretations of the content of the same set of symbols are completely different. If mutual misunderstanding happens, no intended content is transmitted from sender to receiver even when the sender and receiver reach and maintain a signaling equilibrium in

a signaling game. That is to say, successful interactions do not necessarily imply successful communication. As a result, the convention formulated by signaling games is insufficient to explain the linguistic convention of meaning and representation. This leads to mutual misunderstanding because the framework overlooks the semiotic nature of convention. According to Peircean semiotics, convention has dual aspects: The convention of the sign vehicle and the convention of the referential relationship. The signaling game theory provides a partial account of the conventional sign vehicle (or legisign) but neglects the referential aspect. The mutual misunderstanding cases can be better explained through this semiotic understanding.

6.2 Lewis-Skyrms signaling games

Lewis-Skyrms signaling games model signaling interactions between senders and receivers. A simple signaling game includes three elements: A sender, signals and a receiver. The sender can perceive the state of world and send signals to the receiver but is incapable of playing any acts. While the receiver cannot perceive the state of world but can play acts with respect to the received signals. The reason why the sender sends signals when perceives the state and the receiver plays acts when receives signals is that they share a common interest.

The sender's strategy (f_s) is to map states of world onto signals while receiver's (f_r) is to map the signals onto acts so that the states can be mapped onto the acts (F). F is composed by f_s and f_r . F is an informational relationship if sender sends a signal when perceives a state of world while receiver plays a proper act with respect to the state after receiving the signal that both get maximal payoff. Therefore, we say that the signals sent by the sender carry the information of the state of world. That is to say, the content or meaning represented by signals is about the states of world.² Thus, sender-receiver configuration (SRC) provides a framework concerning the fundamental questions of meaning.

David Lewis (1969) formalized the convention of signals as signaling games. Signaling games are a kind of coordination games. Presume a simple signaling game. There are two possible states of world, S_1 and S_2 , two kinds of available signals, M_1 and M_2 , and two kinds of possible acts, A_1 and A_2 . The signals carry no content a priori. That is to say, there are no predetermined relationships between the states and the signals (f_s), and the signals and the acts (f_r). Each act is a proper response to exactly one state. Then, the problem is how the acts can properly respond to the states of world. Therefore, the problem signaling games model is "state-act coordination problem" (Huttegger 2007a), namely, how an informational F is possible.

The solution given by signaling games is the communication through signaling. The information of states is transmitted from senders to receivers. Assume senders and receivers have a common interest,³ if the act being played is proper to a state, both players get the same payoff. If the act is not proper to the state, then none of them gets any payoff. This is shown in Table 6.1, given that the payoff for both players is 1.

Table 6.1 A simple signaling game

	A_1	A_2
S_1	1, 1	0, 0
S_2	0, 0	1, 1

Then, how does communication, namely the combination of a proper f_S and a proper f_R to an informational F , emerge from interactions between senders and receivers in signaling games?

Given the simple signaling game, the sender has four possible strategies, $f_{S1}(S_1, M_1; S_2, M_2)$, $f_{S2}(S_1, M_2; S_2, M_1)$, $f_{S3}(S_1, M_1; S_2, M_1)$ and $f_{S4}(S_1, M_2; S_2, M_2)$. In the cases of f_{S1} and f_{S2} , the sender sends one signal when perceives a state and otherwise sends the other. In the cases of f_{S3} and f_{S4} , the sender always sends the same signal irrespective of what the state of world is. The receiver also has four possible strategies, $f_{R1}(M_1, A_1; M_2, A_2)$, $f_{R2}(M_1, A_2; M_2, A_1)$, $f_{R3}(M_1, A_1; M_2, A_1)$ and $f_{R4}(M_1, A_2; M_2, A_2)$. In the cases of f_{R1} and f_{R2} , the receiver plays one act when receives a signal and plays the other otherwise. In the cases of f_{R3} and f_{R4} , the receiver always plays the same act irrespective of which signal it receives. As F is the combination of f_S and f_R , there are 16 possible combined strategies. Table 6.2 is the payoff matrix of each combined strategy.

Table 6.2 Possible combined strategies in a simple signaling game

	f_{R1}	f_{R2}	f_{R3}	f_{R4}
f_{S1}	1, 1	0, 0	1/2, 1/2	1/2, 1/2
f_{S2}	0, 0	1, 1	1/2, 1/2	1/2, 1/2
f_{S3}	1/2, 1/2	1/2, 1/2	1/2, 1/2	1/2, 1/2
f_{S4}	1/2, 1/2	1/2, 1/2	1/2, 1/2	1/2, 1/2

As we can see, in the cases of $F_1(f_{S1}, f_{R1})$ and $F_2(f_{S2}, f_{R2})$, both players get the maximal payoff. F_1 and F_2 are strict Nash equilibria in the game. A Nash equilibrium has the property that any player who unilaterally deviates her/his strategy from the equilibrium cannot do better. A strict Nash equilibrium has the property that any player who unilaterally deviates his/her strategy from the equilibrium actually does worse. Lewis calls Nash equilibria in signaling games signaling systems (Lewis 1969, 135; Skyrms 1996, 83). It implies that when a signaling game is at a strict Nash equilibrium, a signal being sent by the receiver conveys the information of a state.

Therefore, F_1 and F_2 are two possible signaling systems in the simple signaling game (Figure 6.1).

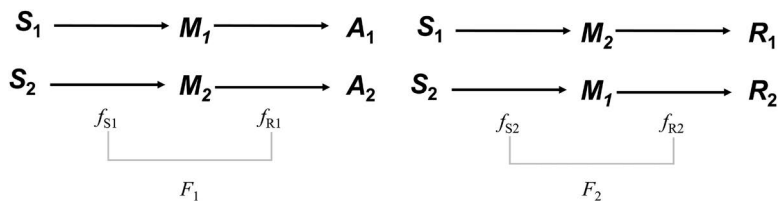


Figure 6.1 Signaling systems in a simple signaling game

Since F_1 and F_2 are two equally effective signaling systems, while only one of them will be actually realized, how do players make the choice between them? Why do they maintain the signaling system being chosen? They are the equilibrium selection and maintenance problems (Skyrms 2004, 50).

For Lewis, who employed rational choice theory to approach signaling games, equilibrium maintenance is not a problem because the structure of signaling games and the strategies of the players are common knowledge. That is to say, both players rationally know that signaling systems are strict Nash equilibria and each of them knows that the other knows she/he knows that they will actually be strictly worse off if any of them unilaterally changes his/her strategy (Lewis 1969, 56). For the equilibrium selection problem, Lewis, following Schelling, appealed to prior agreement, precedent, or salience (Lewis 1969, 158). It means that one of the equally effective signaling systems will stand out or be salient for the players.

However, Skyrms argues that the solutions provided by Lewis cannot explain away the vicious circle pointed out by Quine (1936). First, if there is common knowledge, where does it come from? How do players know that it is common knowledge? Preexisting communication is required to explain common knowledge, while the communication must presume signaling systems with which to convey information. Then, we go back to where we started. Second, the salience of a signaling system is a psychological factor which is not included in the structure of signaling games (Huttegger 2007a). It is *ad hoc* for solving the problem.

Instead, Skyrms argues that interacting individuals spontaneously learn to signal and species spontaneously evolve signaling systems (Skyrms 2010a, 6). It can be explained by evolutionary game theory. Using replicator dynamics to model the simple signaling game and stimulating it on the computer, he finds that one of two signaling system strategies always takes over the whole population. It implies that signaling system strategies are evolutionarily stable strategies in the game. In a signaling game,

[s]trategy I is evolutionarily stable if for all alternative strategies J, either (1) the payoff of I played against I is greater than that of J played against I or (2) I and J have equal payoffs played against I but J has a greater payoff than I when played against J.

(Skyrms 1996, 88)

Therefore, “If just signaling system strategies are evolutionarily stable, then if one strategy takes over the population we should expect it to be a system strategy” (Skyrms

1996, 90). A simple analysis can be made to show that strict Nash equilibria are evolutionarily stable strategies (Skyrms 1996, 90–94). Then, if a strategy takes over the population in an evolutionary signaling game, it must be a signaling system strategy. Furthermore, for any sender-receiver game with the same number of signals as states and strategies, “a strategy is evolutionarily stable if and only if it is a signaling system strategy” (Skyrms 1996, 90). Later, signaling systems are understood as signaling equilibria, which are dynamically stable in terms of dynamic game theory (Huttegger 2007a; Huttegger et al. 2010; Skyrms 2010; Huttegger and Zollman 2011; Huttegger 2014). In the following, I will use signaling equilibria to understand signaling systems.

As we can see, in the simple signaling game, signaling equilibria are strict Nash equilibria that the sender and the receiver get maximal payoff through proper acts. Therefore, we can say that signaling equilibria are success conditions of acts for the sender and the receiver. “In the signaling equilibrium of a Lewis sender-receiver game, information is transmitted from sender to receiver...” (Skyrms 2010, 45). Since a signaling equilibrium is necessary and sufficient for signaling system strategy, the success conditions of acts are the success conditions of communication. If we can explain signaling equilibria in signaling games, then conventionality of meaning is explained.

With the breakthrough, many philosophers believe that theories of meaning have been re-oriented. “[A] theory of meaning is a theory of sender-receiver coordination” (Godfrey-Smith 2012). Some may argue against that Lewis-Skyrms signaling game theory promises a lot while it produces little, or, at least, not enough. What the question the theory attempts to answer is “Can meaning of words arise spontaneously, by chance?” (Skyrms 2010a). It is about the conventionality of language and logic. Notice that the question is not just about how animal signaling spontaneously arises, but about linguistic words.

As we all know, there is a strong tradition in philosophy that there are essential differences between human beings and other species of living organisms. Human beings are rational, having epistemic beliefs, while other species are non-epistemic. Animals communicate information through signaling while human beings use language to express their epistemic belief. With respect, language has been thought of as being essentially different from animal signals. While what signaling game theory models for now are interactions between non-epistemic agents. Can those models be generalized to explain language? Or, can interacting epistemic agents who share no common language spontaneously develop a new language?

It seems that signaling game theory denies the differences between language and other signals in nature. Linguistic symbols are different from animal signals in degree, not in kind.⁴ At the beginning of his masterwork, Skyrms claims that,

In a famous essay on meaning, H. Paul distinguished between natural and non-natural meaning. Natural meaning depends on association arising from natural processes. I say that all meaning is natural meaning.

(Skyrms 2010a, 1)

Thereafter, all meanings can be explained by SRC. “[S]ender-receiver systems are a natural kind, something that nature builds repeatedly in different contexts

and at different scales” (Godfrey-Smith 2013). Language is complex. It is built spontaneously from simple ones. Therefore, in order to explain complex signals like language, we should begin with the simplest ones. Step by step, language can be explained at last. This is what the proponents of the theory actually have done in the past few years. An epitomized work on the topic is Skyrms (2010a).

Some may argue against that there may be no fundamental difference between language and animal signals, but humans are rational, having epistemic beliefs, while animals do not. What Lewis-Skyrms signaling game theory discusses so far is about signaling interactions between non-epistemic players (Skyrms 1996; 2004). Can it be generalized to explain signaling interactions between epistemic players like human beings? Or, can signaling games accommodate signaling interactions between rational players? Does epistemic belief as mental language play any role in signaling games?

The answer to the first and second questions is affirmative. “Accounts framed in terms of mental language, or ideas or intentions can fit perfectly within sender-receiver games” (Skyrms 2010a, 7). However, the answer to the third question is negative. “Within the framework of Lewis signaling games this is not necessary” (Skyrms 2010a, 9). No rationality is needed to explain convention (Huttegger 2014). In other words, epistemic belief plays no role in signaling games. Therefore, it is not necessary to consider epistemic belief of players in signaling games. The claim that converging to and maintaining a signaling equilibrium in a signaling game implies a signaling system is also true for epistemic agents. In the following sections, I will show that this is not without challenging as player’s epistemic beliefs may lead to mutual misunderstanding in which a signaling equilibrium is reached and maintained, but no communication occurs.

6.3 Mutual misunderstanding

Once upon a time in the reign of Chenghua in the Ming Dynasty, China, a Taoist from Japan came to China seeking a challenge to the Masters of China through magical fights. Accidentally, the minister sent by the emperor of China to look for a qualified Master to rise to the challenge misidentified a drunk butcher who slaughters pigs and sells meat as a Master. The Taoist and the butcher had a magical fight on a square in the Forbidden City. The emperor and some ministers watched the fight as audiences. One segment of the fight was for one to guess hand gesture riddles posed by the other.

Hand gesture riddles are a traditional game in China. A simple gesture riddle game consists of two players. One of them poses a riddle to the other through a hand gesture. The other guesses the riddle also by showing a hand gesture in turn, according to her/his interpretation of the meaning of the riddle. It is an interactive process. The process of the fight was like this:

The Taoist showed one finger to the butcher (R_1). In turn, the butcher showed two fingers to the Taoist (R_2). The Taoist showed three fingers to respond to it (R_3). Then, rather than showing four fingers, the butcher showed five fingers (R_4). With respect, the Taoist flapped the top of his head (R_5). The butcher patted his belly

(R_6). Then, the Taoist quit the fight and admitted that he had failed. The butcher claimed that he won. The fight ended.⁵

The emperor witnessed the whole event and was aware of the end but knew nothing about gesture riddles. In order to know what exactly had happened, he asked the Taoist and the butcher separately. The Taoists said that they posed and answered gesture riddles to each other. R_1 means Buddha worship, R_2 means that two saints protect his body, R_3 means that Three Sovereigns run the country, R_4 means that Five Emperors are the supremacy, R_5 means that the Heaven is above his head, and R_6 means that the Buddha sits in his heart.

In contrast, the butcher gave an entirely different explanation. He thought that the Taoists knew that he was a butcher selling meat and thus wanted to bargain with him. R_1 means that the Taoist wanted to buy a pig, R_2 means that there was no problem even if he had wanted to buy two pigs, R_3 means that the pig he wanted to buy should weigh around 30 kg, R_4 means that the weight of the pigs the butcher owned was at least 50 kg, R_5 means that he wanted the pig's head, and R_6 means that the pig offal also belongs to him.

The marvelousness of the story is that those misunderstandings did not prevent but coincidentally promote interactions between the Taoist and the butcher. There are two orders of misunderstanding in the story. The first order is the mutual misunderstanding between the Taoist and the butcher. It means that there was no communication between them, though they interacted with each other successfully. That is to say, no information is transmitted from the butcher to the Taoist, or, *vice versa*, through those gestures. The Taoist misunderstood the butcher that he posed gesture riddles to him, while the butcher misunderstood the Taoist that he bargained with him about his pigs. Each of them made their decision according to their subjective prediction. The form of prediction-decision-making is as follows: Given players 1 and 2 in a game,

(a) 1 maximizes his payoff predicting that 2's decision is governed by (b);

(b) 2 maximizes his payoff predicting that 1's decision is governed by (a).

(Kaneko and Kline 2010)

While the predictions of the Taoist and the butcher were mismatched since there was no common knowledge shared by them. It resulted in mutual misunderstanding.

The second order of misunderstanding is that the emperor misunderstood the interactions between the Taoist and the butcher as a process of meaningful communication according to his perception of the event. He must have believed that there was common knowledge between them, as there were successful interactions and an agreed result. In short, mutual misunderstanding coincidentally promotes to an end that all participants agreed with.

If we formulate the mutual understanding scenario through game theory, the game presented in the scenario is a complex game composed of three games: One in the emperor's mind (g_e), one individually believed by the Taoist (g_t), and one individually imagined by the butcher (g_b).⁶ Each of the participants individually

believed that the structure of the game in his own mind was common knowledge. The end of the story is a Nash equilibrium. That is to say, the interactions between the Taoist and the butcher reached a Nash equilibrium without common knowledge and communication of information. How could this be possible? Suppose the structures of the payoff of g_e , g_i , and g_b are as given in Tables 6.3–6.5⁷ (Kaneko and Kline 2010).

Table 6.3 g_e

	Q_1	Q_2
P_1	5, 5	1, 6
P_2	6, 1	3, 3

Table 6.4 g_i

	Q_1	Q_2
P_1	5, 0	1, 0
P_2	6, 0	3, 3

Table 6.5 g_b

	Q_1	Q_2
P_1	0, 5	0, 6
P_2	0, 1	3, 3

As we can see, although g_e , g_i , and g_b are different games, the Nash equilibria of those games are the same ones, namely (P_2, Q_2) . Let's formalize the structure of the game. Given B_i as the personal belief of a player i , C as common knowledge, the mutual misunderstanding scenario in *Magical Fight* can be formulated as follows:

$$\begin{aligned} & \vdash (B_e C(g_e), B_i C(g_i), B_b C(g_b)) \\ & \rightarrow (B_e C(\text{Nash}(P_2, Q_2)) \wedge B_i C(\text{Nash}(P_2, Q_2)) \wedge B_b C(\text{Nash}(P_2, Q_2))) \end{aligned}$$

In epistemic logic (Kaneko 2002; Kaneko and Suzuki 2002), $\vdash$ means that the entire statement following it is provable, and $\rightarrow$ means that the statement being right of it is concluded from the statement being left of it. The part after $\vdash$ and before $\rightarrow$ consists of assumptions or axioms. Formula (6.1) states that given that the emperor individually believed that the structure of g_e must be common knowledge for the Taoist and the butcher, the Taoist individually believed that the structure of g_i was common knowledge for himself and the butcher, and the butcher individually believed that the structure of g_b was common knowledge for himself and the Taoist, we can conclude that the equilibrium reached at the end is the intersection of Nash equilibria of g_e , g_i , and g_b .

Some may complain that formula (6.1) is misleading as there is no real common knowledge in the story despite those individually believed kinds. Actually, it is easy to revise it into a formula without common knowledge. Given Γ_i as the situation supposed by player i and D_i as the dominant strategy of the game believed by player i , then formula (6.1) can be rewritten as follows:

$$\begin{aligned} &\vdash (B_e(\Gamma_e), B_t(\Gamma_t), B_b(\Gamma_b)) \rightarrow (B_e(\text{Nash}(P_2, Q_2)) \\ &\quad \wedge B_t(D_t(P_2)) \wedge B_b(D_b(Q_2))) \end{aligned}$$

The emperor individually believed that the Taoist and the butcher meaningfully communicated with each other through posing gestures, the Taoist individually believed that the butcher posed gesture riddles to him, and the butcher individually believed that the Taoist bargained on his pigs with him. As a result, the Taoist employed the dominant strategy in his individually supposed game when he interacted with the butcher, and *vice versa*. Coincidentally, the combination of the strategies reaches the Nash equilibrium of the game in the emperor's mind. I call the formula (6.2) a mutual misunderstanding. Then, can we extend mutual misunderstanding to signaling games?

6.4 Mutual misunderstanding in signaling games

It seems obvious that there is no problem in treating the interactions between the Taoist and the butcher through posing gestures to each other as a process of signaling. The Taoist and the butcher played the role of sender and receiver half the time. The Taoist played the role of sender and the butcher of the receiver when he posed a gesture to the butcher, and *vice versa*. Gestures posed by them are signals. All the players and their observers individually believed that the interactions were a process of signaling. However, the same gestures have different meanings for the Taoist and the butcher. The equilibrium was coincidentally reached. No information is transmitted through those gestures. We can say that it is a process of pseudo-signaling. It deceived the emperor who witnessed the magical fight.

As I have illustrated in [Section 6.2](#), given a signaling game, converging to and maintaining signaling equilibrium strategies are the necessary and sufficient conditions for signaling system strategies. If the mutual misunderstanding can be generalized to the framework of the Lewis-Skyrms signaling game, then the claim will be incorrect. That is to say, reaching and maintaining a signaling equilibrium is not sufficient for communicating information from a sender to a receiver. Successful interactions do not imply successful communication. In other words, in a signaling game, it is possible that the sender and receiver mutually misunderstand each other while their interactions still reach and maintain a signaling equilibrium.

However, the application of mutual misunderstanding argument to signaling games is not as straightforward as it looks. Some may argue that we cannot get the conclusion from the argument given in [Section 6.3](#) because signaling games are different in kind from those games in [Section 6.3](#). First, signaling games are state-act coordination games, while the games implied in *Magical Fight* are not.

Second, those games have no equilibrium selection problem. Third, the interests of the butcher and of the Taoist are in conflict, while senders and receivers share a common interest in signaling games. Fourth, mutual misunderstanding is a result of the analysis in terms of rational choice theory rather than evolutionary dynamics.

The best way to respond to the first and second queries is to construct a mutual misunderstanding argument consisting of coordination games. For the third one, the sender and the receiver in a signaling game share complete common interest is ideal. Some researchers have proved that signaling is possible when players share partial or even no common interest (Godfrey-Smith and Martínez 2013; [Martínez and Godfrey-Smith 2016](#)). The fourth query has no effect on the argument because, on the one hand, as I have said above, it is easy to prove that the Nash equilibria in those games are signaling equilibria. On the other hand, rational choice theory and evolutionary dynamics are just different ways to explain how the interactions between players in a game reach and maintain an equilibrium. While what mutual misunderstanding argument doubts is whether convergence to a signaling equilibrium is sufficient for communication. How to explain equilibrium affects the argument a little. Then, the next problem is to construct a mutual misunderstanding argument composed of coordination games.

Lewis-Skyrms signaling games are state-act coordination games that enable the coordination of receiver's acts with states of world. A different kind of coordination game is about coordination of sender's act with receiver's act. It is common that act-to-act coordination is achieved through signaling, such as the example that two men row a boat together coordinating their acts through drum. Act-to-act coordination is different from state-act coordination in that a state is exogenous. It is not chosen by one of the players. Now it is not a state that happens by chance but sender's act. Of course, there is no clear distinction between state-act coordination games and act-to-act coordination games. Many cases have a mixture of both ([Godfrey-Smith 2013](#)). It is obvious that the framework of Lewis-Skyrms signaling games can accommodate act-to-act coordination with no problem.

Suppose there are two players, sender and receiver, playing a coordination game. There are two kinds of possible acts that sender can take, two kinds of signals available, M_1 and M_2 , and two kinds of possible acts that receiver may take. The possibilities that sender takes two acts are equal. When sender takes an act by chance, she/he sends a signal to receiver. After receiving a signal sent by sender, receiver takes an act accordingly. There is no predetermined relationship between sender's acts, signals, and receiver's acts. Different from the simple signaling game introduced in [Section 6.2](#), the sender and receiver here are rational, having some epistemic belief about the game they play. Sender individually believes that the game, g_1 , she/he plays with receiver is one likes stag hunt game, while receiver personally believes that she/he is playing a game like battle of sex, g_2 , with sender. They do not know each other's actual payoff when they interact since there is no common knowledge and priori communication system between them.

Assume a stag hunt game, if a player plays a hare strategy, no matter what strategy the other player plays, she/he always gets 1 unit of payoff. If the player plays a stag strategy, her/his payoff depends on the strategy the other player plays. If it is a hare strategy, then she/he gets 0 units. If stag strategy, then 2 units. This is shown in [Table 6.6](#).

Table 6.6 Stag hunt game, g_1

	Hare (R_H)	Stag (R_S)
Hare (S_H)	1, 1	1, 0
Stag (S_S)	0, 1	2, 2

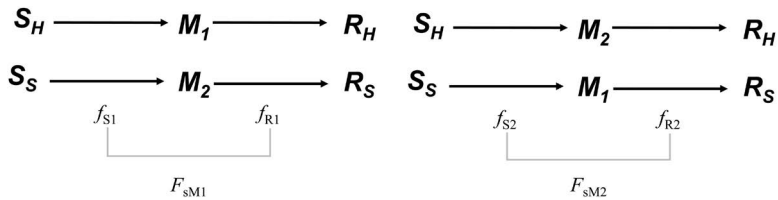
The game in sender's mind is a mixture of state-act coordination game and act-to-act coordination game. Same as the stag hunt game, there are two kinds of prey in the game, hare and stag. Hares are many while stags are rare in the environment. Suppose that hares are always there and stags appear half the time of hares and by chance. Since sender is rational, she/he wants to hunt hare alone when no stag appears and to hunt stag with receiver when stags show up. Therefore, she/he has to develop an information system with receiver in order to get maximal payoff. As there is no common knowledge between her/him, sender has to learn from experience through reinforcement learning, for example.

In the game, when sender sees hares or a stag and hunts it, he always sends a signal to receiver.⁸ When receiver receives a signal and plays her/his hunting strategy with respect. Then, sender has four strategies, $f_{S1}(S_H, M_1; S_S, M_2)$, $f_{S2}(S_H, M_2; S_S, M_1)$, $f_{S3}(S_H, M_1; S_S, M_1)$ and $f_{S4}(S_H, M_2; S_S, M_2)$, and sender thinks that receiver also has four strategies, $f_{R1}(M_1, R_H; M_2, R_S)$, $f_{R2}(M_1, R_S; M_2, R_H)$, $f_{R3}(M_1, R_H; M_2, R_H)$ and $f_{R4}(M_1, R_S; M_2, R_S)$. The structure of the payoff of possible combined strategies in sender's mind is shown in Table 6.7.

Table 6.7 Possible combined strategies in g_1 as a signaling game

	f_{R1}	f_{R2}	f_{R3}	f_{R4}
f_{S1}	1, 1	2, 2	1, 0	0, 1
f_{S2}	1, 0	0, 1	1, 1	2, 2
f_{S3}	1, 1	0, 1	1, 0	2, 2
f_{S4}	1, 0	2, 2	1, 1	0, 1

As we can see, (f_{S1}, f_{R1}) and (f_{S2}, f_{R2}) are two Nash equilibria. Therefore, two possible signaling systems are shown in Figure 6.2.

Figure 6.2 Two possible signaling systems in the g_1

Through sender's own reinforcement learning, one of the signaling systems will be reached and maintained normally.

In contrast, receiver individually believes that the game, g_2 , she/he plays with sender is one like the battle of sexes game. Imagine there is a couple. The husband prefers to watch football during the weekend, while the wife prefers the movie. However, for both of them, the most important thing is to stay together to spend the weekend. Therefore, when they do not stay together, they get no payoff. If they go to watch football, the payoff gotten by the husband is more than the wife. In turn, if they go to watch a movie, the wife will get more payoff. Then, the distribution of the payoff can be assumed as shown in Table 6.8.

Table 6.8 Battle of sexes game, g_2

	Football (H_F)	Movie (H_M)
Football (W_F)	1, 2	0, 0
Movie (W_M)	0, 0	2, 1

Receiver subjectively imagines that when sender decides to do an activity, she/he always sends a signal to the other one. When receiver receives a signal, she/he decides what she/he will do. Then, receiver personally believes that sender has four strategies, $f_{S1}(W_F, M_1; W_M, M_2)$, $f_{S2}(W_F, M_2; W_M, M_1)$, $f_{S3}(W_F, M_1; W_M, M_1)$ and $f_{S4}(W_F, M_2; W_M, M_2)$, and she/he has four strategies with respect, $f_{R1}(M_1, H_F; M_2, H_M)$, $f_{R2}(M_1, H_M; M_2, H_F)$, $f_{R3}(M_1, H_F; M_2, H_F)$ and $f_{R4}(M_1, H_M; M_2, H_M)$. The structure of the payoff of possible combined strategies in receiver's mind is shown in Table 6.9.

Table 6.9 Possible combined strategies in g_2 as a signaling game

	f_{R1}		f_{R2}		f_{R3}		f_{R4}	
f_{S1}	1, 2		0, 0		1, 2		0, 0	
		2, 1		0, 0		0, 0		2, 1
f_{S2}	0, 0		1, 2		1, 2		0, 0	
		0, 0		2, 1		0, 0		2, 1
f_{S3}	1, 2		0, 0		1, 2		0, 0	
		0, 0		2, 1		0, 0		2, 1
f_{S4}	0, 0		1, 2		1, 2		0, 0	
		2, 1		0, 0		0, 0		2, 1

As we can see, (f_{S1}, f_{R1}) and (f_{S2}, f_{R2}) are two Nash equilibria. Therefore, two possible signaling systems are shown in Figure 6.3.

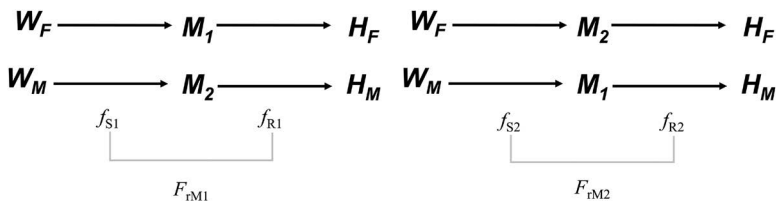


Figure 6.3 Two possible signaling systems in the g_2

Normally, one of the signaling systems will be converged to and maintained through receiver's reinforcement learning.

With the epistemic beliefs in mind, sender and receiver interact with each other through signaling. The strategies sender has are $f_{S1}(S_H, M_1; S_S, M_2)$, $f_{S2}(S_H, M_2; S_S, M_1)$, $f_{S3}(S_H, M_1; S_S, M_1)$ and $f_{S4}(S_H, M_2; S_S, M_2)$. While the strategies receiver has are $f_{R1}(M_1, H_F; M_2, H_M)$, $f_{R2}(M_1, H_M; M_2, H_F)$, $f_{R3}(M_1, H_F; M_2, H_F)$ and $f_{R4}(M_1, H_M; M_2, H_M)$. Coincidentally, H_F is proper to S_H that sender gets 1 unit payoff and receiver gets 2 units; H_M is proper to S_S that sender gets 2 units and receiver gets 1 unit. It is necessary to note that it is not the states sender perceives are matters for receiver but sender's acts. This is shown in Table 6.10.

Table 6.10 The actual payoff of acts coordination

	H_F	H_M
S_H	1, 2	1, 0
S_S	0, 0	2, 1

When sender plays S_H , sender thinks that receiver plays either R_H or R_S . While given sender plays S_H , if receiver gets 2 units when playing H_F , then she/he believes that sender plays W_F ; if she/he gets no payoff when playing H_M , then sender plays W_F . When sender plays S_S , if no payoff is gotten, then she/he thinks that receiver plays R_H ; if she/he gets 2 units, then receiver plays R_S . Given sender plays S_S , if receiver gets no payoff when playing H_F , then she/he believes that sender plays W_M ; if 1 unit is gotten when playing H_M , then sender plays W_M . Accordingly, the structure of the actual payoff of possible combined strategies is given in Table 6.11.

Table 6.11 Actual possible combined strategies in the mutual misunderstanding scenario

	f_{R1}	f_{R2}	f_{R3}	f_{R4}
f_{S1}	1, 2	1, 0	1, 2	1, 0
f_{S2}	2, 1	1, 2	0, 1	0, 0
f_{S3}	1, 0	0, 0	2, 1	0, 0
f_{S4}	1, 2	1, 0	1, 2	1, 0
	0, 0	2, 1	0, 0	2, 1
	1, 0	1, 2	1, 2	1, 0
	2, 1	0, 0	0, 0	2, 1

There are two possible signaling equilibria in the matrix (Figure 6.4).

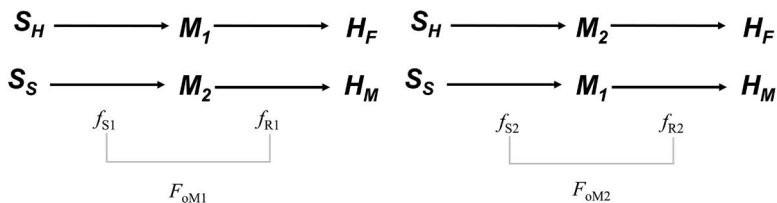


Figure 6.4 Two actual possible signaling equilibria in the g_2

As we can see, mutual misunderstanding does not impede but promotes convergence to signaling equilibrium. In turn, it is obvious that there is no information transmitted when two mutually misunderstand each other.

Using the formal definition of mutual misunderstanding introduced in the last section, we can formalize the scenario as such:

Given B_i as the personal belief of a player i , C as the personally believed common knowledge, the situation can be formalized as,

$$\begin{aligned} \vdash & \left(B_1 C(g_1), B_2 C(g_2) \right) \rightarrow \left(B_1 C \left(\text{Nash}(f_{s1}, f_{r1}) \right) \wedge B_2 C \left(\text{Nash}(f_{s1}, f_{r1}) \right) \right) \\ & \vee \left(B_1 C \left(\text{Nash}(f_{s2}, f_{r2}) \right) \wedge B_2 C \left(\text{Nash}(f_{s2}, f_{r2}) \right) \right) \end{aligned}$$

Given Γ_i as the situation supposed by player i and D_i as the dominant strategy of the game believed by player i , then formula (6.3) can be rewritten as follows:

$$\begin{aligned} \vdash & \left(B_1(\Gamma_1), B_2(\Gamma_2) \right) \rightarrow \left(\wedge B_1(D_1(f_{s1})) \vee B_2(D_2(f_{r1})) \right) \\ & \wedge \left(B_1(D_1(f_{s2})) \wedge B_2(D_2(f_{r2})) \right) \end{aligned}$$

As we can see, formula (6.4) shows that mutual misunderstanding, namely a game can reach and maintain an equilibrium without information exchanges, is still possible given that the games are subjectively interpreted as coordination games by participants while no intended information is transmitted from sender to receiver. Next, let us analyze the contents of the mutual misunderstanding scenario.

6.5 Objective and subjective content in mutual misunderstanding

When agents in signaling games are rational, there is subjective and objective content, as Skyrms argues. A signal's subjective content may not align with the signal's objective content at all (Skyrms 2010a; 2010b). However, we will not follow Skyrms' theory but take an account developed by Shea et al. (2018). We know that no theory on content is widely accepted and every theory is controversial (Birch 2014; Scarantino 2015; Skyrms and Barrett 2019; Isaac 2019). We intend to put aside these issues since this is not our concern.⁹ Instead, we simply use the account of Shea et al. (2018) as a tool to show that the subjective content diverges from the objective content in the scenario of mutual misunderstanding.

Signals carry meaning or content of some kind. According to Shea et al. (2018), signals in signaling games carry two kinds of content: Informational and functional. Informational content is about the world's states/acts. Information is the change made by the signal to the probability of those states/acts. The informational content of a signal M_j is the vector of the binary logarithm of post-signal

probabilities of the states, $\log_2[P(S_i|M_j)]$, in which P is a probability and S_i is a state in question.¹⁰ Then, the informational content of M_j about the states S_i of the world is as follows:

$$\langle \log_2[P(S_1 | M_j)], \log_2 [P(S_2 | M_j)], \dots, \log_2 [P(S_i | M_j)], \dots \rangle$$

The informational content of M_j about acts A_i is:

$$\langle \log_2[P(A_1 | M_j)], \log_2[P(A_2 | M_j)], \dots, \log_2[P(A_i | M_j)], \dots \rangle.$$

As we can see, the informational content of a signal is the distribution of the probabilities of the world states/acts conditional on that signal, in which at least some of the probabilities of the states/acts are different from the unconditional probabilities of the states/acts. The informational content of a signal is determined by the unconditional probabilities of the states/acts and the selection of the signal and the act (the sender's and receiver's strategies).

The informational content of a signal may not have anything to do with the equilibrium of a signaling game, namely, signaling systems. What matters is functional content. Functional content is about stabilizing the signaling systems and the beneficial outcome of sender-receiver coordination. The signals in a signaling game have functional content only if some selection process maintains a signaling system in the game (Shea et al. 2018, 1015–1016). Roughly speaking, functional content is a vector representing each state's relative importance to stabilizing a signaling system. Unlike informational content, which is measured by probabilities, the functional content of a signal, for each state, is measured by the degree of involvement of that state, pairing with a receiver's specific act through that signal, in stabilizing a signaling system. Nevertheless, "[t]he two kinds of content have the same form—distributions over states of the world, one reflecting posterior probabilities and one reflecting functional involvement" (Shea et al. 2018, 1017). There are also subjective contents when agents have beliefs. The sender has subjective content over the receiver's strategies, and the receiver has subjective content over the sender's strategies, and both have subjective content over the states (Skyrms 2010b, 162). With the conception of content in hand, let us analyze the content in the scenario of mutual misunderstanding.

In a mutual misunderstanding scenario, there are three coordination signaling games: g_1 , g_2 , and the played game. g_1 lives in the sender's belief, while g_2 lives in the receiver's. Slightly differing from the conception of content introduced above, there is informational content about the states of the world and acts.¹¹ Informational content about acts is separate from that of the sender's acts and the receiver's acts. Functional content is about the relative importance of the combination of each state of the world pairing with the sender's act and the receiver's act to the stabilization of a signaling system. Now, suppose the signaling system F_{oM1} is actually reached

and stabilized. What will be the objective and subjective content of signals? Objective content lives in the realized game. Table 6.12 shows the objective content in vector form.

Table 6.12 The relationship between informational and functional content in actual game

		<i>Informational content about the states of the world</i>	<i>Informational content about the sender's acts</i>	<i>Informational content about the receiver's acts</i>	<i>Functional content</i>
Signal	M_1	$\langle 1, 0 \rangle$; Hare	$\langle 1, 0 \rangle$; S_H	$\langle 1, 0 \rangle$; H_F	$\langle 1, 0 \rangle$; Hare- S_H - H_F
	M_2	$\langle 1, 0 \rangle$; Stag	$\langle 0, 1 \rangle$; S_s	$\langle 0, 1 \rangle$; H_M	$\langle 0, 1 \rangle$; Stag- S_s - H_M

In F_{oM1} , signal M_1 makes the state of Hare certain and completely rules out the state of Stag, so post-signal probabilities of the states of the world are $\langle 1, 0 \rangle$. This is the informational content about the states of the world. In our case, it pools with the informational content about the sender's acts. So, they share the same distribution. The distinction between informational content about the states of the world and about the sender's acts presented here is conceptually rather than structurally in the game. M_1 makes act H_F certain and completely rules out act H_M , so post-signal probabilities of the receiver's acts are $\langle 1, 0 \rangle$. The distribution of the probabilities of the sender's and receiver's possible acts constitutes the informational content about the acts regarding M_1 . The functional content of M_1 is the functional involvement of the combination of Hare, S_H with H_F , short for Hare- S_H - H_F , in the stabilization of F_{oM1} . Hare- S_H - H_F takes F_{oM1} to be stabilized and contributes nothing to the stabilization of F_{oM2} , so the functional act of M_1 is $\langle 1, 0 \rangle$. Because F_{oM1} , an equilibrium in the game, is reached and maintained, the informational content about the states of the world, the sender's acts and the receiver's acts, and the functional content are the same. M_1 is sent only when the sender perceives a Hare and takes S_H , the receiver only takes H_F when receiving M_1 . The stabilization of F_{oM1} is achieved if and only if the sender perceives Hare and takes S_H when the receiver takes H_F .

The subjective content of the sender is different from the objective content given in Table 6.12. Table 6.13 shows the sender's subjective content in vector form.

Table 6.13 The relationship between informational and functional content for the sender

		<i>Informational content about the states of the world</i>	<i>Informational content about the receiver's acts</i>	<i>Functional content</i>
Signal	M_1	$\langle 1, 0 \rangle$; Hare	$\langle 1, 0 \rangle$; R_H	$\langle 1, 0 \rangle$; Hare- S_H - R_H
	M_2	$\langle 1, 0 \rangle$; Stag	$\langle 0, 1 \rangle$; R_s	$\langle 0, 1 \rangle$; Stag- S_s - R_s

The signal M_1 makes Hare certain and completely rules out Stag, so the post possibilities of the states of the world is $\langle 1, 0 \rangle$. This is what happened, so the sender's informational content about the world's states is in line with that of the objective content. As objective content, the distinction between those two kinds of content is conceptual, not structurally, in the game. Since the sender is always certain about the act she takes, there is no information about her acts for her. The sender's informational content about the receiver's acts differs from that of the objective content. The informational content about the receiver's acts of M_1 in the actual game is $\{\langle 1, 0 \rangle; H_F\}$ while that in g_1 is $\{\langle 1, 0 \rangle; R_H\}$. Her functional content, $\{\langle 1, 0 \rangle; \text{Hare-}S_H\text{-}R_H\}$, is also different from the objective functional content which is $\{\langle 1, 0 \rangle; \text{Hare-}S_H\text{-}H_F\}$. Note that the difference is not in quantity but in kind. The sender is wrong because she does not know what is in the receiver's mind. The sender and the receiver have no common knowledge after all. She can only rely on her own beliefs.

Same as the sender, the receiver's subjective content differs from objective content. It also differs from the sender's subjective content. Table 6.14 shows the receiver's subjective content in vector form.

Table 6.14 The relationship between informational and functional content for the receiver

		<i>Informational content about the sender's acts</i>	<i>Functional content</i>
Signal	M_1	$\langle 1, 0 \rangle; W_F$	$\langle 1, 0 \rangle; W_F\text{-}H_F$
	M_2	$\langle 0, 1 \rangle; W_M$	$\langle 0, 1 \rangle; W_M\text{-}H_M$

There is no informational content about the states of the world for the recipient because she is incapable of perceiving. Same as the sender, the receiver is always certain about the acts taken by herself, so there is no information about the acts she took for her. Compared with the objective content of M_1 , the informational content about the sender's acts and functional content for the receiver are different. She is wrong about the content of the sender's acts and functional content because she has no way to know the sender's mind and can only appeal to her own beliefs. The subjective content of M_1 for the receiver is different from that for the sender. It results from the absence of shared common knowledge between the sender and receiver. They are isolated in their subjective worlds by themselves as long as the payoff is optimal for both sides.

The analysis corresponds neatly to what Skyrms has argued in the paragraph, though we approach content in signaling games with different terms.

These subjective probabilities are just degrees of belief; they may not align with the objective probabilities. Then, each signal carries two additional kinds of subjective information. There is subjective information about how the receiver will react, which lives in the sender's degrees of belief. This interests a sender who wants to get a receiver to do something. There is

subjective information about what state the sender observed, which lives in the receiver's degrees of belief. This is of interest to a receiver who wants to use the sender as a source of information about the states. Both sender and receiver use these kinds of information in decision-making. Both sender and receiver strive (1) to act optimally given their subjective probabilities and (2) to learn to bring subjective probabilities in concordance with the objective probabilities in the world. They may or may not succeed.

(Skyrms 2010a, 44–45)

What is newly revealed in the mutual misunderstanding scenario is that signals' subjective and objective content may differ not only in quantity but also in kind. It implies that the object content can be independent of the subjective content in a sender-receiver signaling game. It does not mean that subjective content cannot be the same as objective content. Of course, they can. The complete mutual misunderstanding introduced here is an idealized model. It is as rare as complete mutual understanding. Then, why signaling game theory results in mutual misunderstanding? A reason may be that it overlooks the semiotic nature of conventionality and only tells a partial story.

6.6 Semiotics of conventionality

Recall the classifications of signs according to the presentational and representational conditions in Peircean semiotics. We can identify two types of arbitrariness and conventionality: Conventional sign vehicles (legisigns) and conventional referential relationships (symbols). When a sign is interpreted as a legisign, it means that the intrinsic properties of the sign qua object have no bearing on its role as the ground of the sign. The intrinsic properties of the sign qua object do not form the basis for its representative character (Liszka 1996, 35–36). When we consider a sign as a legisign, we are concerned only with the relationship between the intrinsic properties of the sign and the properties that serve as the grounds for it being a sign. Since a sign, as a legisign, is defined by its presentational properties, it can serve not only as a symbol but also as an icon or index. For example, the symbol :) is a legisign but used as an icon of a smiling face due to its resemblance to one's smiling; a directional arrow on a road lane is a legisign serving as an index of a left-turn lane due to the spatial continuity of the direction it points to.

From the representational condition, the type of sign that is conventional is the symbol. That is, the representational relationship between a symbol and its object does not depend on the intrinsic properties of the sign vehicle and that of object. For example, the relationship between the word "dog" and the animal dog has no connection to the presentational properties of the word or the animal. According to this definition, since qualisigns and sinnsigns are defined by their intrinsic properties, only legisigns—whose intrinsic properties are irrelevant to their status as signs—can serve as symbols. Therefore, symbols exhibit dual aspects of

conventionality: That of the sign vehicle (legisign) and that of the referential relationship (symbolic reference). Philosophical traditions in language often overlook this semiotic distinction.

This raises the question: Which aspect of conventionality does evolutionary signaling game theory explain? Or is there a unified theory that accounts for both aspects? To explore this, let's revisit the theory from the perspective of semiotics.

An essential difference between evolutionary game theory and semiotics is that the former considers reference at the population level, defining it as intersubjective (or as a “proper function”, or “stabilizing function” of signs, in Millikan's terms), while the latter focuses on the individual level, treating reference as the result of interpretation. As I argued in [Chapter 4](#), a problem with this evolutionary approach is that it assumes the past determines the proper function of a sign, yet a sign typically refers to the present. What natural selection and alike mechanisms provide are constraints that limit the ways a sign can be interpreted.¹² If these constraints fulfill their proper function, they must contribute to the self-maintenance and self-generation of teleodynamics. Despite these difficulties, the minimal competence required by the approach—natural selection and trial-and-error learning—implies semiotic interpretation. Specifically, in the case of natural selection, it must involve operational interpretation contributing to the self-maintenance and self-generation of teleodynamics. In signaling games, this contribution is represented and measured by payoffs. In the case of trial-and-error learning, interpretation must involve higher-order forms of interpretation. Let's explore why.

A prerequisite for players to converge to and maintain a signaling equilibrium is that they must have at least interpretive competence to recognize the type of signal and distinguish it from other types. Recall the sensitive autogen model discussed in [Chapter 4](#), which has an adaptive advantage over the prototype with no sensitive surface. When the substrates are rich, their surface is more easily disrupted as more substrates bond to it. Thus, it has indexical competence, with the bonding point serving as an indicator of the substrate-rich environment. To have this competence, it must be able to distinguish between the necessary substrates and others. Operational interpretation performs this function by bonding the needed substrates while not attaching to others. Accordingly, the bonding point is an indexical sinnsign. In this sense, it is not conventional. However, it can become conventional.

Imagine another version of the autogen whose surface is not sensitive to the substrates directly involved in the reciprocal catalysis. Instead, its surface is sensitive to the product of a specific chemical reaction. This type of chemical reaction has several products, and one of which is the substrates needed by the reciprocal catalysis inside the autogen. Since the bonded products and the needed products are produced by the same reactions, they aggregate together. Thus, when the surface bonds enough of these products and breaks up, the autogen also opens up to the needed substrates. The bonding point remains an indicator of a substrate-rich environment, but it is also possible for the surface to bond with other types of products. Although the substrates to the surface bonds are highly constrained by chemical laws, the autogen can use different chemical bonds as indicators of the same environment.

We can frame this scenario within the sender-receiver framework of signaling games. The sender is the type of chemical reaction producing various products, including the substrates needed by the autogen. The possible signals are the products that can be bonded by the surface. The receiver is the autogen. The possible world states are a substrate-rich environment and a substrate-barren environment. The possible acts are opening up or not. We can imagine that a type of surface bonding with a specific product could be stabilized by natural selection. In this case, it becomes conventional and arbitrary. Accordingly, it is a legisign because its intrinsic properties do not form the basis of the representational relationship. As I have argued, the referential function of the bonding point is indexical. Therefore, the bonding point is an indexical legisign of a substrate-rich environment. Thus, what evolutionary game theory and related approaches, like Millikan's biosemantics, explain is the conventionality of sign vehicles, not of representational relationships. This distinction can also be observed in animal signals.

Seyfarth et al. (1980a; 1980b) famously showed that vervet monkeys have three distinct alarm calls for three types of predators: An eagle in the sky, a leopard and a snake on the ground. This is a widely cited example in signaling game studies. The specific call used for each predator is not determined a priori but by the dynamics of sender-receiver interactions. Therefore, the alarm calls are conventional. However, this conventionality pertains to the sign vehicle, not the referential relationship. A call is spatial-temporally contiguous with the appearance of a predator. Therefore, the alarm calls serve as indexes. Accordingly, they are indexical legisigns. Without distinguishing between the conventionality of sign vehicles and that of referential relationships, the Lewis-Skyrms sender-receiver signaling game theory cannot account for mutual misunderstandings. Such cases, however, can be explained through the concept of the regrounding of symbolic reference, as discussed in [Chapter 5](#).

As we have argued, different types of referential relationships are hierarchically structured. Symbolic representation asymmetrically depends on indexical representation, which in turn depends on iconic representation. Conversely, when a symbol cannot be interpreted as a symbol, it may be interpreted as an index or even an icon—a descent down a hierarchy of diminishing interpretive competence (Deacon 1997, 72–73). This is clearly illustrated in the story of *the Magical Fight*. When the butcher saw the gestures made by the Taoist, he interpreted them as symbols because they resemble gestures he used in his own past. This is iconic interpretation. However, he could not interpret them symbolically because he lacked knowledge of the Taoist's symbol network. What he interpreted was that these gestures indicate the Taoist's intentions. He must figure out their meaning. This is similar to translating a language he has never encountered before. To translate a language without a dictionary, we must reground it within linguistic practice to learn the indexical relationships between symbols, and between symbols and objects. In the case of the magical fight, the butcher only has his personal experience as a butcher to draw on. To understand the gestures, he regrounded them with his own experience. He iconically found that the gestures are similar to those he used when bargaining with customers. He treated them as bargaining gestures and thus

established indexical relationships between them and positions them within his own bargaining symbol network. Since these gestures are grounded in the practice of bargaining, the indexical relationships between the bargaining gestures and their relevant objects in the situation are transitive to the Taoist's gestures. Accordingly, the butcher symbolically exchanged gestures with the Taoist. The Taoist also re-grounded the butcher's gestures within his own symbol network.

As we can see, these interpretive processes occur at the individual level, in the minds of the butcher and the Taoist. This aspect is excluded by the evolutionary signaling games theory, so it cannot be explained by that framework. This is an exception to the signaling game theory, but it is a common feature in our everyday linguistic practice. While the Lewis-Skyrms framework offers a useful and insightful understanding of meaning, it only provides part of the picture. Through semiotics, we gain a more comprehensive understanding of meaning, intentionality, reference, or representation.

6.7 Conclusion

In the preceding sections, I have argued that there is a flaw in the Lewis-Skyrms signaling game theory, using the mutual misunderstanding argument. In signaling games, where information is communicated from senders to receivers, signaling systems are identified with the equilibria of the games. The mutual misunderstanding argument suggests that merely converging to and maintaining an equilibrium in a signaling system is insufficient to guarantee successful communication. Mutual misunderstanding can still occur. This argument challenges signaling game theory because the theory overlooks the distinction between two types of conventionality: That of the sign vehicle and that of referential relationships. Convention is not only a phenomenon of behavioral interactions but is, at its core, a semiotic phenomenon. What evolutionary game theory explains is the conventionality of the sign vehicle—the constraints placed on how people interpret signals. With the distinction between these two types of conventionality, along with the symbol regrounding account, we gain a more comprehensive understanding of the conventionality of symbolic reference. This allows for a better explanation of cases involving animal signals and mutual misunderstanding.

Notes

- 1 [Millikan \(2005, 2\)](#) strongly argues that it is a fallacy to treat language convention a type of coordination convention because it requires neither coordination, regular conformity, nor rational underpinning. Instead, she develops a Darwinian evolution by natural selection-inspired account. Nevertheless, we will see that the core elements of her account, reproduction and weight of precedent, are integrated into Skyrms' evolutionarily dynamic reformulation of signaling game and thus are compatible with signaling game theory.
- 2 The content of signals in signaling games is a complex topic of signaling game study. There is a whole chapter, [Chapter 3](#), in [Skyrms \(2010a; 2010b\)](#) that discusses the concept

of information in signaling games. The content of signals in the simple signaling games is not just about the states of the world but about acts. Millikan names this kind of signals as pushimi-pullyus (Millikan 2004, 77–78). Even in the simple signaling games, we should distinguish informational content, which is about the states of the world, from functional content, which is obtained only when the game is at a signaling equilibrium and maintained by selection process (Shea, Godfrey-Smith and Cao 2018)

- 3 Common interest between sender and receiver is not necessary for signaling games to reach and maintain an equilibrium, partial common interest, or even complete conflict interest, is enough, though not as stable as common interest cases, as Godfrey-Smith and Martínez (2013) and Martínez and Godfrey-Smith (2016) have shown.
- 4 Some may argue that Ruth Millikan (1989a) has clearly argued that animal representations are different from human beliefs in at least six fundamental different ways. Since Skyrms (2010, 1) and others (Godfrey-Smith 2012, for instance) claim that the framework of signaling games is in the same way with Millikan's biosemantics, it implies that they would agree with Millikan's claims on the subject. If this is true, Millikan thinks that, compared with human beliefs, animal representations lack of differentiation between indicative and imperative representations and inference. While these are Lewis-Skyrms signaling games tries to explain. (Skyrms 2004, Chapter 4; Huttegger 2007b) It means that the explanations provided by signaling games have overstepped animal signals. Then, my argument still holds as they apply the same framework to explain non-epistemic and epistemic signaling interactions.
- 5 The description of the story is an excerpt of a famous Chinese monologue comic talk given by Baorui Liu called *Magical Fight*.
- 6 Different from Kaneko and Kline's original understanding, which claims that g_e is an objective, general and abstractive game in contrast with those two personally believed games, I identify g_e as the game which was personally believed by the emperor, who was one of the audience of the magical fight. In the Japanese version of the story, the audience is set up as an outside observer who objectively knows all the information of the game. While in *Magical Fight*, as I analyzed above, the emperor misunderstood the interactions as a process of communication, while there was no information being transmitted between the two players. Of course, we will see that there is a game objectively, truthfully and generally representing the real structure of the story. In Kaneko and Kline (2010), Kaneko and Kline introduce a Japanese version of the story, *Konnyaku Mondô*, and discuss the implications for game theory. In addition, two friends of mine told me a Jewish and a British version of the story.
- 7 Rigorously, the games introduced here do not fit the story very well. The purpose here is to give an instance of mutual misunderstanding.
- 8 Some may argue against that there is no motivation for sender to send signals when she/he hunts hare because she/he always gets fixed payoff no matter what the other player does. This is true. However, no signal sent is also a kind of signal as an adage says, "No news is good news!"
- 9 People who may have an interest in the debate of the content of signals in signaling games can see Birch (2014), Shea et al. (2018), Skyrms and Barrett (2019), and Godfrey-Smith (2020).
- 10 Another way to measure the informational content given by Skyrms (2010a; 2010b) is that it is the vector of the binary logarithm of the ratio of the conditional to the unconditional probability of that state/act. This may be a better way. But, as Shea et al. (2018) argue, the way employed in the paper makes some formal comparison between informational and functional content possible. In addition, in order to solve the problem of misinformation, Birch (2014) develops a theory that a signal would carry the content is at the nearest separating equilibrium of the underlying evolutionary dynamics.

- 11 When discussing functional content, [Shea et al. \(2018\)](#) only mention the functional involvement of the states of the world. We think that it is an ellipsis because they note at the end of [Section 3.1](#) that, “As Skyrms notes, a signal can carry information about both the states of the world perceived by the sender and about acts produced by the receiver. Here, we will only discuss informational content about the state of the world.” Both states and acts matter for the functional content. In addition, since two of the signaling games we discuss in the present paper are a mixture of state-act and act-to-act coordination games, and we pool the states of the world with the sender’s acts, the informational content about the sender’s and receiver’s acts matters.
- 12 This result has implications for the content analysis made in [Section 6.5](#). There is no objective content but constraints. Therefore, what is measured by the part of objective content is actually constraints produced by the interactions.

References

- Adraans, Pieter. (2010). A Critical Analysis of Floridi's Theory of Semantic Information. *Knowledge Technology Policy*, 23: 41–56.
- Allen, Colin and Jacob Neal. (2020). Teleological Notions in Biology. *The Stanford Encyclopedia of Philosophy* Edward N. Zalta (ed), <https://plato.stanford.edu/archives/spr2020/entries/teleology-biology/>
- Barbieri, Marcello. (2007). Is the Cell a Semiotic System? In Marcello Barbieri (ed), *Introduction to Biosemiotics: The New Biological Synthesis*. Dordrecht: Springer, pp. 179–207.
- Barbieri, Marcello. (2019). Code Biology, Peircean Biosemiotics, and Rosen's Relational Biology. *Biological Theory*, 14: 21–29.
- Bar-Hillel, Yehoshua. (1953/1955). Information and Content. *Synthese*, 9(3/5): 299–305.
- Bar-Hillel, Yehoshua. (1955). An Examination of Information Theory. *Philosophy of Science*, 22(2): 86–105.
- Bar-Hillel, Yehoshua and Rudolf Carnap. (1953a). An Outline of a Theory of Semantic Information. In: Reprinted in Y. Bar-Hillel (ed), *Language and Information: Selected Essays on Their Theory and Application*. 1964, Reading, MA: Addison-Wesley, pp. 221–274.
- Bar-Hillel, Yehoshua and Rudolf Carnap. (1953b). Semantic Information. *The British Journal for the Philosophy of Science*, 4(14): 147–157.
- Barrett, Jeffrey A. (2009). The Evolution of Coding in Signaling Games. *Theory and Decision*, 67(2): 223–237.
- Barrett, Jeffrey A. (2016). On the Evolution of Truth. *Erkenntnis*, 81(6): 1323–1332.
- Barrett, Jeffrey A., C. Cohran and Brian Skyrms. (2020). Hierarchical Models for the Evolution of Compositional Language. *Philosophy of Science*, 87(5): 910–920.
- Barrett, Jeffrey A. and T. LaCroix (2020). Epistemology and the Structure of Language. *Erkenntnis*, 87: 953–967. <https://doi.org/10.1007/s10670-020-00225-4>
- Bates, Marcia J. (2005). Information and Knowledge: An Evolutionary Framework for Information Science. *Information Research: An International Electronic Journal*, 10(4): 1–24.
- Bates, Marcia J. (2006). Fundamental Forms of Information. *Journal of the American Society for Information Science and Technology*, 57(8): 1033–1045.
- Bateson, G. (1972). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Chicago, Illinois: University of Chicago Press.
- Bateson, G. (1979). *Mind and Nature: A Necessary Unity*. Hampton: Hampton Press.
- Bedau, Mark A. (2008). What is life? In Sahorta Sarkar and Plutynski Anya (eds), *Companion to the Philosophy of Biology*. Oxford: Blackwell, pp. 455–471.
- Berger, C. R. and R. J. Calabrese. (1975). Some Exploration in Initial Interaction and Beyond: Toward a Developmental Theory of Communication. *Human Communication Research*, 1: 99–112.
- Bergstrom, C and M. Rosvall. (2011). The Transmission Sense of Information. *Biology and Philosophy*, 26(2): 159–176.

- Birch, Jonathan. (2014). Propositional Content in Signalling Systems. *Philosophical Studies*, 171: 493–512.
- Brenner, Joseph. (2014). Information: a Personal Synthesis. *Information*, 5: 134–170.
- Brier, S. (2008). *Cybersemiotics: Why Information Is Not Enough*. Toronto, Ontario: University of Toronto Press.
- Burgin, Mark. (2010). *Theory of Information: Fundamentality, Diversity and Unification*. Singapore: World Scientific.
- Capurro, Rafael. (2009). The Past, Present and Future of the Concept of Information. *Triple C*, 7(2): 125–141.
- Capurro, Rafael and Birger Hjøland. (2003). The concept of information. In B. Cronin (ed.), *Annual Review of Information Science and Technology* (ARIST), 37 (Chapter 8), 343–411.
- Capurro, Rafael and Birger Hjøland. (2005). The Concept of Information. *Annual Review of Information Science and Technology*, 37(1): 343–411.
- Carnap, Rudolf and Yehoshua Bar-Hillel. (1952a). An outline of a theory of semantic information, Technical Report No. 247, *Research Laboratory of Electronics*, Massachusetts Institute of Technology.
- Carnap, Rudolf and Yehoshua Bar-Hillel. (1952b). Semantic Information. *The British Journal for the Philosophy of Science*, 4(14): 147–157.
- Cashman, T. (2008). What connects the map to the territory. In M. Barbieri and J. Hoffmeyer (eds), *A Legacy for Living Systems: Gregory Bateson as Precursor to Biosemiotics*. New York: Springer, pp. 45–58.
- Chaitin, G. J. (1987). *Algorithmic Information Theory*. New York: Cambridge University Press.
- Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Chemero, A. (2003). An Outline of a Theory of Affordances. *Ecological Psychology*, 15(2): 181–195.
- Chomsky, Noam. (1968). *Language and Mind*. San Diego: Harcourt Brace.
- Chomsky, Noam. (1980). *Rules and Representations*. New York: Columbia University Press.
- Christie, Joshua R., Carl Brusse, Pierrick Bourrat, Peter Takacs and Paul E. Griffiths (2022). Are Biological Traits Explained by Their ‘Selected Effect’ Functions? *Australasian Philosophical Review*, 6(4): 335–359. <https://doi.org/10.1080/24740500.2024.2370630>.
- Collier, J. (1996). Information Originates in Symmetry Breaking. *Symmetry: Culture & Science*, 7: 247–256.
- Collier, John. (2003). Hierarchical Dynamical Information Systems With a Focus on Biology. *Entropy*, 5: 102.
- Crick, F. (1958). On Protein Synthesis. *Symposia of the Society for Experimental Biology*, 12: 138–163.
- Cubitt, R. and R. Sugden. (2003). Common Knowledge, Salience and Convention: A Reconstruction of David Lewis’ Game Theory. *Economics and Philosophy*, 19: 175–210.
- Cummins, Robert. (1975). Functional Analysis. *The Journal of Philosophy*, 72(20): 741–765.
- Davidson, Donald. (1987). Knowing One’s Own Mind. *Proceedings and Addresses of the American Philosophical Association*, 60: 441–458.
- Deacon, Terrence. (1997). *The Symbolic Species: The Coevolution of Language and Brain*. New York: W. W. Norton & Company.
- Deacon, Terrence. (2003). Universal grammar and semiotic constraints. In M. Christiansen and S. Kirby (eds), *Language Evolution*, Oxford: Oxford University Press, pp. 111–139.
- Deacon, Terrence. (2006). Reciprocal Linkage between Self-Organizing Processes Is Sufficient for Self-Reproduction and Evolvability. *Theoretical Biology*, 1(2): 136–149.
- Deacon, Terrence. (2007). Shannon-Boltzmann-Darwin: Redefining Information. Part 1. *Cognitive Semiotics*, 2: 123–148.
- Deacon, Terrence. (2008). Shannon-Boltzmann-Darwin: Redefining Information. Part 2. *Cognitive Semiotics*, 2: 167–194.

- Deacon, Terrence. (2012a). *Incomplete Nature: How Mind Emerged from Matter*. New York: W. W. Nordon & Company.
- Deacon, Terrence. (2012b). The symbolic concept. In Kathleen R. Gibson and Maggie Tallerman (eds), *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press, pp. 393–405.
- Deacon, Terrence. (2017). Information and reference. In Gordana Dodig-Crnkovic and Raffaella Giovagnoli (eds), *Representation of Reality: Humans, Other Living Organism and Intelligent Machines*. Heidelberg: Springer, pp. 3–15.
- Deacon, Terrence. (2018). Beneath symbols: Convention as a semiotic phenomenon. In Steven C. Hayes and David Wilson (eds), *Evolution and Contextual Behavioral Science: An Integrated Framework for Understanding, Predicting, and Influencing Human Behavior*. Oakland, CA: Context Press, pp. 79–98.
- Deacon, Terrence. (2020). Teleodynamics: Specifying the dynamical principles of intrinsically end-directed processes. *Proceedings of the IASAE 2020: International Conference on Thermodynamics 2.0*, ICT2.0: 2020-W1xx.
- Deacon, Terrence. (2021). How a Molecule Became a Sign? *Biosemiotics*, 14: 537–559.
- Deacon Terrence, Srivastava A and JA Bacigalupi. (2014). The Transition from Constraint to Regulation at the Origin of Life. *Front Biosci* (Landmark Ed). 19(6):945–57. <https://doi.org/10.2741/4259>.
- DE Tienne, André. (2005). Information in Formation: A Peircean Approach. *Cognitio*, 6(2): 149–165.
- Dodig-Crnkovic, G. (2011). Dynamics of Information as Natural Computation. *Information*, 2(3): 460–477.
- Dretske, Fred. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: The MIT Press.
- Dretske, Fred. (1986a). Misrepresentation. In Radu Bogdan (ed), *Belief: Form, Content, and Function*. New York: Oxford University Press, pp. 17–36.
- Dretske, Fred. (1986b). *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: The MIT Press.
- Duzi, M. (2010). The Paradox of Inference and the Non-Triviality of Analytic Information. *Journal of Philosophical Logic*, 39(5): 473–510.
- Eigen, M. and P. Schuster. (1979). *The Hypercycle: A Principle of Natural Self-Organization*. Berlin: Springer-Verlag.
- Fallis, Don. (2015). What Is Disinformation? *Library Trends*, 63(3): 401–426.
- Ferguson, Thomas. (2015). Two Paradoxes of Semantic Information. *Synthese*, 192: 3719–3730.
- Floridi, Luciano. (2004). Outline of a Theory of Strongly Semantic Information. *Minds and Machines*, 14(2): 197–222.
- Floridi, Luciano. (2010). *Information: A Very Short Introduction*. Oxford: Oxford University Press.
- Floridi, Luciano. (2011). *Philosophy of Information*. Oxford: Oxford University Press.
- Fodor, Jerry. (1984). Semantics Wisconsin Style. *Synthese*, 59: 231–250.
- Fodor, Jerry. (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.
- Fox, C. J. (1983). *Information and Misinformation: An Investigation of the Notions of Information, Misinformation, Informing, and Misinforming*. Westport, CT: Greenwood Press.
- Franke, M. and E. O. Wagner. (2014). Game Theory and the Evolution of Meaning. *Language and Linguistic Compass*, 8(9): 359–372.
- Frege, Gottlob. (1892/1948). Sense and Reference. *The Philosophical Review*, 57(3): 209–230.
- Fresco, Nir, Patrick McGivern and Aditya Ghose. (2017). Information, Veridicality, and Inferential Knowledge. *American Philosophical Quarterly*, 54(1): 61–75.
- Fresco, Nir and Michaelis Michael. (2016). Information and Veridicality: Information Processing and the Bar-Hillel/Carnap Paradox. *Philosophy of Science*, 83: 131–151.

- Gánti, Tibor. (1979/2003). *Chemoton Theory: Theory of Living Systems*. New York: Kluwer Academic/Plenum Publisher.
- García-Valdecasas, Miguel. (2022). On the Naturalisation of Teleology: Self-Organisation, Autopoiesis and Teleodynamics. *Adaptive Behavior*, 30(2): 103–117.
- García-Valdecasas, Miguel and Terrence Deacon. (2024). Biological Functions Are Causes, Not Effects: A Critique of Selected Effects Theories. *Studies in History and Philosophy of Science*, 103: 20–28.
- Godfrey-Smith, Peter. (1996). Spence and Dewey on Life and Mind. In M. Biden (ed), *The Philosophy of Artificial Life*. Oxford: Oxford University Press, pp. 314–331.
- Godfrey-Smith, Peter. (2003). Genes Do Not Encode Information for Phenotypic Traits. In C. Hitchcock (ed), *Contemporary Debates in Philosophy of Science*. Blackwell: London.
- Godfrey-Smith, Peter. (2007). Conditions for Evolution by Natural Selection. *Journal of Philosophy*, 104(2007): 489–516.
- Godfrey-Smith, Peter. (2012). Brian Skyrms's Signals: Evolution, Learning, and Information. *Mind*, 120: 1288–1297.
- Godfrey-Smith, Peter. (2013). What do Generalizations of Lewis Signaling Model Tell us about Information and Meaning. *A lecture given at MIT and Columbia*, November 2013.
- Godfrey-Smith, Peter. (2014). Sender-Receiver Systems Within and between Organisms. *Philosophy of Science*, 81: 866–878.
- Godfrey-Smith, Peter. (2016). Individuality, Subjectivity and Minimal Cognition. *Biology and Philosophy*, 31: 775–796.
- Godfrey-Smith, Peter. (2020). In the Beginning There Was Information? *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 80: 1–7.
- Godfrey-Smith, Peter and M. Martínez. (2013). Communication and Common Interest. *PLoS Computational Biology*, 9: e1003282.
- Godfrey-Smith, Peter and Kim Sterelny, “Biological Information”, The Stanford Encyclopedia of Philosophy (Summer 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2016/entries/information-biological/>.
- Gould, S. J. and E. S. Vrba. (1982). Exaptation- a Missing Term in the Science of Form. *Paleobiology*, 8: 4–15.
- Gregesen, Niels. (2010). God, Matter, and information: Towards a stoicizing logos christology. In Paul Davies and Niels Gregesen (eds), *Information and the Nature of Reality: From Physics to Metaphysics*. Cambridge: Cambridge University Press, pp. 319–348.
- Grice, Paul. (1957). Meaning. *Philosophical Review*, 66: 377–388.
- Griffiths, Paul. (2001). Darwinism and Developmental Systems. In S. Oyama, P. Griffiths and R. Gray (eds), *Cycles of Contingency: Developmental Systems and Evolution*. Cambridge, MA: MIT Press, pp. 195–218.
- Griffiths, Paul. (2001). Genetic Information: A Metaphor in Search of a Theory. *Philosophy of Science*, 68: 394–412.
- Harms, William. (2004). Primitive Content, Translation, and the Emergence of Meaning in Animal Communication. In D. Kimbrough and U. Griebel (eds), *Evolution of Communication Systems: A Comparative Approach*. Cambridge, MA: MIT Press, pp. 31–48.
- Harms, William. (2006). What Is Information? Three Concepts. *Biological Theory*, 1(3): 230–242.
- Harnad, Stevan. (1990). The Symbol Grounding Problem. *Physica D*, 42: 335–346.
- Heras-Escribano, M. (2020). The Evolutionary Role of Affordances: Ecological Psychology, Niche Construction, and Natural Selection. *Biology and Philosophy*, 35: 30.
- Hoffmeyer, Jesper. (2008). *Biosemiotics: An Examination into the Signs of Life and the Life of Signs*. Scranton: University of Scranton Press.
- Hoffmeyer, Jesper and Clause Emmeche. (1991). Code-duality and the Semiotics of Nature. In Anderson Myrdene and Merrell Floyd (eds), *On Semiotic Modeling*, Berlin, Boston: Mouton De Gruyter, pp. 117–166.

- Hofkirchner, W. (2013). *Emergent Information: A Unified Information Framework*. Singapore: World Scientific.
- Huttegger, S. (2014). How Much Rationality Do We Need to Explain Conventions. *Philosophy Compass*, 9(1): 11–21.
- Huttegger, S., B. Skyrms and R. Smead. (2010). Evolutionary Dynamics of Lewis Signaling Games: Signaling Systems vs. Partial Pooling. *Synthese*, 172: 177–191.
- Huttegger, S., B. Skyrms, P. Tarrès and E. Wagner. (2014). Some Dynamics of Signaling Games. *Proceedings of the National Academy of Sciences*, 111(Supplement 3): 10873–10880.
- Huttegger, Simon. (2007a). Evolution and the Explanation of Meaning. *Philosophy of Science*, 74: 1–27.
- Huttegger, Simon. (2007b). Evolutionary Explanations of Indicatives and Imperatives. *Erkenntnis*, 66: 409–436.
- Huttegger, S.M. and K. J. Zollman (2011). Signaling games - Dynamics of Evolution and Learning. In A. Benz, C. Ebert, G. Jäger and R. van Rooij (eds), *Language, Games, and Evolution. Lecture Notes in Computer Science*, vol 6207. Berlin, Heidelberg: Springer, pp. 160–176.
- Isaac, A. (2019). The Semantic Latent in Shannon Information. *British Journal for the Philosophy of Science*, 70: 103–125.
- Jablonka, E. (2002). Information: Its Interpretation, Its Inheritance, and Its Sharing. *Philosophy of Science*, 69(4): 578–605.
- Kaneko, M. (2002). Epistemic Logics and Game Theory: Introduction. *Economic Theory*, 9: 7–62.
- Kaneko, M. (2005). *Game Theory and Mutual Misunderstanding: Scientific Dialogues in Five Acts*. Berlin: Springer.
- Kaneko, M. and N-Y. Suzuki. (2002). Bounded Interpersonal Inferences and Decision Making. *Economic Theory*, 9: 63–104.
- Kaneko, Mamoru and J. Kline. (2010). Two Dialogues on Epistemic Logic and Inductive Game Theory. *Advances in Mathematical Research*, 12: 1–44.
- Kauffman, Stuart. (2000). *Investigations*. Oxford: Oxford University Press.
- Kauffman, Stuart K. (1993). *The Origin of Order: Self-Organization and Selection in Evolution*. New York: Oxford University Press.
- Kolmogorov, Andrea N. (1963). On Tables of Random Numbers. *Sankhya: The Indian Journal of Statistics, Series A*, 25(4): 369–376.
- Kolmogorov, Andrea N. (1968). Three Approaches to the Quantitative Definition of Information. *International Journal of Computer Mathematics*, 2(1–4): 157–168.
- Korzybski, Alfred. (1933). Science and Sanity: An introduction to non-Aristotelian systems and general semantics. *The International Non-Aristotelian Library Pub. Co.* pp. 747–761.
- Kullback, S. and R. Leibler. (1951). On Information and Sufficiency. *Annals of Mathematical Statistics*, 22: 79–86.
- Landauer, R. (1961). Irreversibility and Heat Generation in the Computing Process. *IBM Journal of Research and Development*, 5: 183–191.
- Levy, A. (2011). Information in Biology: A Fictionalist Account. *Nous*, 45: 640–657.
- Lewis, David. (1969). *Convention: A Philosophical Study*. Cambridge: Harvard University Press.
- Lewontin, R. (1970). The Units of Selection. *Annual Review of Ecology and Systematics*, 1: 1–18.
- Li, Ming and Paul Vitányi. (2008). *An Introduction to Kolmogorov Complexity and Its Applications*, 3rd ed. New York: Springer.
- Linski, Wolfgang. (2010). Information: A Conceptual Investigation. *Information*, 1: 74–118.
- Liszka, James Jakob. (2016). How Signs Convey Information: A Peircean Approach. *Chinese Semiotic Studies*, 12(1): 45–66.
- Liszka, James Jakob. (1996). *A General Introduction to the Semiotic of Charles Sanders Peirce*. Bloomington and Indianapolis: Indiana University Press.

- Lombardi, Olimpia, Federico Holik and Leonardo Vanni. (2016). What Is Shannon Information? *Synthese*, 193: 1983–2012.
- Long, Bruce. (2014). Information Is Intrinsically Semantic but Alethically Neutral. *Synthese*, 191: 3447–3467.
- Luhmann, Niklas. (1984). *Soziale Systeme*. Frankfurt am Main: Suhrkamp.
- Mackay, D. M. (1953). Generators of Information. *Proceedings of the Second Symposium on Applications of Communication Theory*, London.
- Margolis, Eric and Stephen Laurence. (2019). Concepts. *The Stanford Encyclopedia of Philosophy*. Edward N. Zalta (ed), URL = <https://plato.stanford.edu/archives/sum2019/entries/concepts/>.
- Maroney, Owen. (2009). Information Processing and Thermodynamic Entropy”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed), URL = <https://plato.stanford.edu/archives/fall2009/entries/information-entropy/>.
- Martínez, M. and P. Godfrey-Smith. (2016). Common Interest and Signaling Games: A Dynamic Analysis. *Philosophy of Science*, 83: 371–392.
- Maturana, Humberto and Francisco Varela. (1972/1980). *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht: D. Reidel Publishing Company.
- Maynard-Smith, John. (2000). The Concept of Information in Biology. *Philosophy of Science*, 67(2): 177–194.
- Millikan, Ruth. (1984). *Language, Thought and Other Biological Categories*. Cambridge: MIT Press.
- Millikan, Ruth. (1989a). Biosemiotics. *Journal of Philosophy*, 86(6), 281–297.
- Millikan, Ruth. (1989b). In Defense of Proper Functions. *Philosophy of Science*, 56(2): 288–302.
- Millikan, Ruth. (1990). Biosemiotics. *The Journal of Philosophy*, LXXXVI(6): 281–297.
- Millikan, Ruth. (2004). *Varieties of Meaning*. Cambridge, MA: The MIT Press.
- Millikan, Ruth. (2005). *Language: A Biological Model*. Oxford: The Oxford University Press.
- Millikan, Ruth. (2017). *Beyond Concepts: Unicepts, Language and Natural Information*. Oxford: Oxford University Press.
- Neander, Karen. (1991a). Function as Selected Effects: The Conceptual Analysis Defense. *Philosophy of Science*, 58: 168–184.
- Neander, Karen. (1991b). The Teleological Notion of a Function. *Australasian Journal of Philosophy*, 69: 454–468.
- Neander, Karen. (1995). Misrepresenting and Malfunctioning. *Philosophical Studies*, 79: 109–141.
- Neander, Karen. (2017). *A Mark of the Mental: In Defense of Informational Teleosemantics*. Cambridge, MA: MIT Press.
- Nöth, Winfrid. (2013). Charles S. Peirce’s Theory of Information: A Theory of the Growth of Symbols and of Knowledge. *Cybernetics and Human Knowing*, 19(1–2): 137–161.
- Parsons, Talcott. (1951). *The Social System*. Glencoe, IL: Free Press.
- Peirce, Charles Sanders. (1931–1935). *The Collected Papers of Charles Sanders Peirce*. C. Hartshorne and P. Weiss (eds), Cambridge, MA: Harvard University Press.
- Peirce, Charles Sanders. ([1893–1913] 1998). *The Essential Peirce: Selected Philosophical Writings*. Peirce Edition Project. Bloomington and Indianapolis: Indiana University Press, vol. 2.
- Peters, John Durham. (1988). Information: Notes Towards a Critical History. *Journal of Communication Inquiry*, 12: 9–23.
- Pinker, Steven. (1994). *The Language Instinct*. New York: Harper Perennial Modern Classics.
- Queiroz, João, Clause Emmeche and Charbel Niño ElHani. (2010). Information and semiosis in living systems: A semiotic approach. In Favareau Donald (ed), *Essential Readings in Biosemiotics: Anthology and Commentary*. Dordrecht: Springer.
- Quine, W. V. (1936). Truth by Convention. In O. H. Lee (ed), *Philosophical Essays for A. N. Whitehead*. New York: Longmans, pp. 90–124.
- Raczaszek-Leonardi, Joana and Terrence Deacon. Symbol ungrounding. (Forthcoming).

- Rasmussen, Steen and Mark Bedau, et al. *Procell: Bridging Nonliving and Living Matter*. Cambridge, MA: The MIT Press, 2009.
- Roche, William and Tomoji Shogenji. (2018). Information and Inaccuracy. *British Journal for the Philosophy of Science*, 69: 577–604.
- Sarkar, Sahotra. (1996a). Biological information: A skeptical look at some Central dogmas of molecular biology. In *The Philosophy and History of Molecular Biology: New Perspectives*. Dordrecht: Kluwer.
- Sarkar, Sahotra. (1996b). Decoding “coding”: Information and DNA. *BioScience*, 46: 857–864.
- Sarkar, Sahotra. (2000). Information in Genetics and Developmental Biology: Comments on Maynard Smith. *Philosophy of Science*, 67(2): 208–213.
- Scarantino, Andrea. (2015). Information as Probabilistic Difference Maker. *Australasian Journal of Philosophy*, 93(3): 419–443.
- Schmidhuber, J. (1997). A computer Scientist’s view of life, the universe, and everything. In Freksa C. (ed), *Foundations of Computer Science: Potential - Theory – Cognition*. Lecture Notes in Computer Science. Springer, pp. 201–208.
- Schulte, Peter and Karen Neander. (2022). Teleological Theories of Mental Content”, *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed), URL = <https://plato.stanford.edu/archives/sum2022/entries/content-teleological/>.
- Searl, J. (1979). What Is an Intentional State? *Mind*, 88(349): 74–92.
- Seyfarth, R. M., D. L. Cheney and P. Marler. (1980a). Monkey Responses to Three Different Alarm Calls: Classification and Semantic Communication. *Science*, 210: 801–803.
- Seyfarth, R. M., D. L. Cheney and P. Marler. (1980b). Vervet Monkey Alarm Calls: Semantic Communication in a Free-Ranging Primate. *Animal Behaviour*, 28: 1070–1094.
- Shannon, Claude. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3): 379–423.
- Shannon, Claude. (1993). *Collected Papers*. In N. J. A. Sloane and A. D. Wyner (ed). New York: IEEE Press.
- Shannon, Claude and Wallen Weaver. (1964). *The Mathematical Theory of Information*. Urbana: The University of Illinois Press.
- Shea, Nicholas. (2007). Representation in the Genome and in Other Inheritance Systems. *Biology and Philosophy*, 22: 313–331.
- Shea, Nicholas. (2018). *Representation in Cognitive Science*. Oxford: Oxford University Press.
- Shea, Nicholas, Peter Godfrey-Smith and Rosa Cao. (2018). Content in Simple Signalling Systems. *British Journal for the Philosophy of Science*, 69: 1009–1035.
- Short, Thomas L. (2007). *Peirce’s Theory of Signs*. Cambridge: The Cambridge University Press.
- Simon, H.A. (1972). Theories of Bounded Rationality. In McGuire, C. B. and R. Radner (eds), *Decision and Organization*. Amsterdam: North-Holland Publishing Co.
- Skyrms, Brian. (1996). *Evolution of the Social Contract*. New York: Cambridge University Press.
- Skyrms, Brian. (2004). *The Stag Hunt and the Evolution of Social Structure*. New York: Cambridge University Press.
- Skyrms, Brian. (2010a). *Signal: Information, Learning and Evolution*. Oxford: Oxford University Press.
- Skyrms, Brian. (2010b). The Flow of Information in Signaling Games. *Philosophical Studies*, 147: 155–165.
- Skyrms, Brian and J. Barrett. (2019). Propositional Content in Signals. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 74: 34–39.
- Stampe, Dennis. (1977). Toward a Causal Theory of Linguistic Representation. In Peter A. French, Theodore E. Uehling Jr., and Howard K. Wettstein (eds), *Midwest Studies in*

- Philosophy: Studies in the Philosophy of Language*. Minneapolis: University of Minnesota Press, vol. 2, pp. 81–102.
- Stegmann, Ulrich E. (2015). Prospects for Probabilistic Theories of Natural Information. *Erkenntnis*, 80: 869–893.
- Steinert-Threlkeld, S. (2016). Compositional Signaling in a Complex World. *Journal of Logic, Language, and Information*, 25(3–4): 379–397.
- Steinert-Threlkeld, S. (2019). Towards the Emergence of Non-Trivial Compositionality. *Philosophy of Science*, 87(5): 897–909.
- Sterelny, K., K. C. Smith, and M. Dickison. (1996). The Extended Replicator. *Biology and Philosophy*, 11: 377–403.
- Stonier, T. (1997). *Information and Meaning: An Evolutionary Perspective*. New York: Springer-Verlag.
- Szilard, L., (1929). On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings. *Zeitschrift für Physik*, 53: 840–856. English translation in *The Collected Works of Leo Szilard: Scientific Papers*, B. T. Feld and G. Weiss Szilard (eds), Cambridge, MA: MIT Press, 1972, pp. 103–129.
- Vanderstraeten, Raf. (2002). Parsons, Luhmann and the Theorem of Double Contingency. *Journal of Classical Sociology*, 2(1): 77–92.
- Varela, Francisco. (1997). Patterns of Life: Intervening Identity and Cognition. *Brain and Cognition*, 34: 72–87.
- von Neumann, John. (1966). *Theory of Self-Reproducing Automata*. Urbana-Champaign: University of Illinois Press.
- Weaver, W. (1949). Recent Contribution to the Mathematical Theory of Communication. In Shannon C. and W. Weaver. *The Mathematical Theory of Communication*. Urbana: The University of Illinois Press. 1964.
- Weber, Andreas and Francisco Varela. (2002). Life after Kant: Natural Purposes and the Autopoietic Foundations of Biological Individuality. *Phenomenology and the Cognitive Science*, 1: 97–125.
- Wheeler, J. (1989). Information, Physics, Quantum: The Search for Links. In W. Zurek (ed), *Complexity, Entropy, and the Physics of Information*. Redwood City, CA: Addison-Wesley.
- Wiener, Nobert. (1948/1961). *Cybernetics, or Control and Communication in the Animal and the Machine*, 2nd ed. New York: The MIT Press/Wiley.
- Wittgenstein, Ludwig. (1921). *Tractatus Logico-Philosophicus*. London: Kegan Paul.
- Wolfram, Stephen. (2002). *A New Kind of Science*. Champaign: Wolfram Media, Inc.
- Wright, Larry. (1973). Functions. *The Philosophical Review*, 82(2): 139–168.
- Zuse, K. (1967). Rechnender Raum. *Elektronische Datenverarbeitung*, 8: 336–344.

Index

Note: *Italicized* page references refer to figures and page references with “n” refer to endnotes.

aboutness 1, 26, 38, 61, 67; aspect of information 4, 6–7, 32; naturalistic approach to 2; normative aspect of 55; *see also* [naturalize intentionality/aboutness](#)

Adraans, Pieter 11

algorithmic theory 1

Aristotle 42–43, 46

artifacts 15, 61–62

artificial intelligence (AI) 90–91

Austin 44

autogenesis 21, 68–71, 75–77, 79;
Deacon’s concept of 60; self-maintaining/self-producing 75

Bar-Hillel, Yehoshua 6–8, 24, 34, 39n1

Bar-Hillel and Carnap’s theory 6–8,
10–11, 13

Bar-Hillel-Carnap Paradox 8, 10–11, 13,
40n8

Bates, Marcia J. 27, 35, 38

Bateson, Gregory 35, 45, 47, 57

biosemiotics 67

Birch, Jonathan 103, 124n10

Boltzmann entropy 49, 51, 53–54

The Boy Who Cried Wolf (Aesop) 86, 91, 93

Burgin, Mark 37–38

Carnap, Rudolf 6–8, 10–11, 13, 30, 34;
see also [Bar-Hillel and Carnap’s theory](#)

causal-historical chain 91, 98

causal relations 14–15, 39–40n3, 88

Chaitin, G. J. 31

Chomsky, Noam 92, 94

code-duality 67, 71–75

code fallacy 86, 101

Collier, John 4

communication 50, 52, 56–57, 64, 66,
84, 92; engineering 1, 23; idea of
information as 47; and information 25;
linguistic 25; mathematical theory of 1,
3–6, 31, 33, 39n1

communication systems 4, 5, 26, 42

complexity/algorithm information 31

comprehensive theory of information 3

conceptual analysis 23–39; information as
tripartite phenomenon 24–28; normative
information 35–38; overview 23–24;
referential information 32–34; structural
information 28–32

condition: interpretative 65, 75;
presentative 65; representative 65, 74;
triadic 65, 75, 87

constraint 30–31, 46–49, 51–57, 59–60,
70–71, 74–80, 86–87, 121, 123

containment 68–71, 76–77, 81n4

content-determinacy challenge 40n4

content indeterminacy problem 55–57, 61

convention 102–123; Lewis-Skyrms
signaling games model 104–108; mutual
misunderstanding 108–111; mutual
misunderstanding in signaling games
111–116; objective/subjective content
in mutual misunderstanding 116–120;
overview 102–104; semiotics of
conventionality 120–123

conventional language 86

conventional meaning 14–20, 28, 33, 86, 91
Convention: A Philosophical Study (Lewis)
102

correlational information 18, 41–42, 48–49,
51–52, 57

Critique of Pure Reason (Kant) 85

Cummins, Robert 64; functional analysis theory 64

Darwinian approach 68

Deacon, Terrence 20–21, 27–28, 31, 35, 38, 43, 45–46, 51–52, 63, 66–69, 71, 75, 78–80, 83, 89, 93–95; concept of autogenesis 21, 60; dynamics theory and biosemiotics 20; dynamic theory of physicality of information 55; interpretation of Peircean semiotics 92; toy model of autogenesis 68

digital codes 74

disinformation 34, 40n6

Dretske, Fred 8, 16–19, 34, 39–40n3, 41, 56, 81n1, 103; context-free theory 17; formulation of natural information 16; informational semantics 15–16, 42; semantic theory of information 8–10, 15–17

dynamics theory and biosemiotics 20

dynamic theory of physicality of information 55

emergent dynamics theory 67

Emmeche, Clause 66, 71, 73, 74

engineering conception of information 2

epistemological indeterminacy objection 62

Floridi, Luciano 7–8, 10–11, 27, 29, 34, 38, 40n7, 40n8, 43, 45; definition of data 45; formulation of the open problems 14; measure of quantity of semantic information 22n1; theory of informational content 11–12; theory of strongly semantic information 10–12

Fodor, Jerry 40n4

form 29–30, 41–49, 57, 66–67, 75, 79, 84, 86, 90, 92

functional analysis theory 64

fundamentalism 42–43, 55–57

genes 68

genetic information 71–75

Grice, Paul 2, 14, 16, 25, 41, 55

Harnad, Stevan 90–91

Hoffmeyer, Jesper 66, 71, 73, 74

“How Molecules Became Signs?” (Deacon) 68

The Human Use of Human Beings:

Cybernetics and Society (Wiener) 42

icon 21, 67, 88–90, 92, 96, 120, 122

iconic relationship 100

index 21, 67, 76, 88–90, 92, 96, 98, 120, 122

information: amount of 4–8, 10; comprehensive theory of 3; concept of 2; in context of communication system 4; defined 4–5, 6, 23; engineering conception of 2; grounding meaning in 1–21; information and teleosemantics 16–20; natural information to conventional meaning 14–20; overview 1–3; physical nature of 2; quantity of 5, 7; semiotic understanding of 64–68; theories 6–14; vector theories of informational content 12–14

informational content 8–10, 49, 117–119, 124n10; 125n11; defined 8, 10; quantity of 10, 11; of signal 117; theories of 6–14; vector theories of 12–14

informational entropy 5, 49

informational relationships 2, 3, 13–14, 29

informational teleosemantics 42

information as tripartite phenomenon 24–28

information theory 3–4, 6, 7, 11, 24, 28, 50, 66

informative signals 12

infosigns 17–18, 41

intentionality 59–62, 64, 102, 123; challenge of 2; naturalistic approach to 3, 15, 20, 41, 48; naturalistic theory of 15; normative aspect of 60; semiotic understanding of 64–68

intentional signs 25

interpretation 59–80; autogenesis 68–71; code-duality 71–75; from constraint to information 75–80; from constraint to semiosis 75–80; genetic information 71–75; operational interpretation 71–75; origin of normativity 68–71; overview 59–60; semiotic understanding of information/intentionality 64–68; symbolic 21, 89–90, 93–94, 96, 101; teleosemantics 60–64

interpretative condition 65, 75

Kant, Immanuel 85

Kolmogorov, Andrey 23, 31; algorithmic theory 1; complexity or algorithm information 31

Kripke, Saul 91, 98

Landaaur principle 40n9

language 82–94, 96, 98–103, 107–108, 121;
 biological model of 84; conventional
 86; dyadic model of 83; informational/
 semantic 73; map metaphor of 83;
 philosophy of 33, 39, 88, 92
 legisign 87–88, 91–98, 100, 104,
 120–122

Lewis, David 38, 102, 104–106

linguistic communication 25; *see also*
 communication

linguistic referential relationships 82

linguistic semantic-mapping function 85

linguistic symbols 1, 28, 83, 88–89, 91,
 107

Liszka, James Jakob 65, 87

Luhmann, Niklas 37

map metaphor 83–87

mathematical mapping relationships 98

meaning 23–24, 27–28, 42–44, 48, 57,
 62–67, 70, 71, 82, 86, 91, 93–94, 96,
 98–99, 102–104, 107, 123; conventional
 14–20, 33; grounding 1–21; natural 2,
 14–16, 25, 28, 55; non-natural 2, 14–15,
 18, 25, 55; normative aspect of 18;
 polysemantic concept 33; of sign 17,
 38, 60

measure of quantity of semantic
 information 22n1

mental states and linguistic symbols 1

metabolism 68

metaphysical problem of information 42

Millikan, Ruth 9, 16–18, 25, 36, 38, 41–42,
 43, 45, 48, 57, 60, 63, 64, 66, 84, 123n1;
 biosemantics 42, 48; semantic-mapping
 function theory 85; theory encounters 48

miscommunication 37

misinformation 34, 40n6, 55–57

molecular biology 71–73

Moses populi informatory 44

multidimensional vector coordinate
 system 98

mutual information 5–6; defined 5;
 quantity of 6

mutual misunderstanding 21, 103–104,
 108–111, 122, 123; objective content in
 116–120; in signaling games 111–116;
 subjective content in 116–120

Nash equilibrium 105, 110–111

natural information 14–20

naturalistic philosophy 19

naturalize intentionality/aboutness 41

natural meaning 2, 25

natural signs 25

natural systems of representation 16

Neander, Karen 2, 18

negative logic 46, 48, 57, 77, 80

non-natural meaning 2, 14, 25

normative information 35–38

normativity, origin of 68–71

operational interpretation 71–75

pan-informationalism 24, 29–30

Parsons, Talcott 37

Peirce, Charles Sanders 19–21, 22n3,
 38, 45, 47, 57, 60, 64–67, 73, 75, 77,
 81n2, 83, 87, 88, 91–92, 95, 101,
 104; definition of a sign 65; idea of
 information 47; semiotic framework
 92; semiotics 3, 20–21, 60, 64–65, 67,
 73, 77, 92, 101, 104, 120; semiotics and
 teleosemantics 65–66; theory of signs/
 semiotics 19–20, 22n3; thinking on
 information 66

philosophy of information 14

physical (negatively) information 41–57;
 content indeterminacy problem 55–57;
 entropies 48–50; formal properties
 48–50; grounding re-presentation on
 51–55; misinformation problem 55–57;
 overview 41–43; place of form in nature
 43–48

physical relationship 29–30, 53

place of form in nature 43–48

Plato 28

Platonic realism 43

presentative condition 65

probability-based approach 10

proper function 61; of biological trait 62;
 defined 60–61; *Normal* conditions
 63–64

qualisign 87, 88

quantity of information 7

quantity of mutual information 6

Quine, W.V. 102, 106

reciprocal catalysis and self-assembly
 68–69

reducible physicalism 42, 43

reference 82–101; hierarchy of 87–90;
 map metaphor 83–87; overview
 82–83; symbolic 89, 90–92; symbol
 regrounding 96–100; symbol
 ungrounding 92–96

- referential information 32–34
- representation 14–16, 19, 20, 43, 51–52, 55–56, 59–62, 64, 75–77, 79–80, 82, 93, 97, 100
- representative condition 65, 74
- Scarantino, Andrea 12–13
- Searle, J. 2
- Second Law of Thermodynamics 49–50, 58n5
- self-contradictory sentence 7–8, 11, 13
- semantic theory of information 8–10, 15–17
- semiotics 3, 20–21, 60, 64–65, 66, 67, 73, 77, 92, 101, 104, 120; biosemiotics theory 67; of conventionality 120–123; interpretation of Peircean semiotics 92
- sender-receiver configuration (SRC) 104
- Shannon, Claude 7–8, 11–13, 20, 22n2, 23, 26, 29, 31–33, 39, 39n1, 42, 43, 48–51, 53–54; communication system by 4; communication theory 33; formal theory 4; general formulas 13; information theory 6, 7; mathematical theory of communication 1, 3–6, 31, 43; theory of communication and thermodynamics 43
- Shannon entropy 49, 53–54
- Shea, Nicholas 18, 116, 124n10; 125n11
- signal 12; defined 12–13; informational content of 117
- signaling equilibria 107, 112, 115, 115
- signaling game theory 21, 38, 104, 107, 108, 120–123, 123n1
- sinnsigns 87–88, 97–99, 120–121
- Skyrms, Brian 12, 18, 34, 38, 102, 106, 108, 124n10
- sophisticated objection 62
- Stampe, Dennis 14, 15, 19, 42; causal theory of representation 14–15, 16, 42
- structural information 28–32, 80
- swampman objection 62
- symbol-grounding problem 67
- “The Symbol Grounding Problem” (Harnad) 90
- symbolic interpretation 21, 89–90, 93–94, 96, 101
- symbolic reference 89, 90–92
- symbols 16; regrounding 96–100; ungrounding 92–96
- teleodynamics 121
- teleological system 68, 71
- teleosemantics 16–20, 60–64, 80
- theorem of double contingency in communication 37
- theory of informational content 11–12
- theory of strongly semantic information 10–12
- thermodynamic entropy 49–52
- triadic condition 65, 75, 87
- vector theories of informational content 12–14
- veridicality 34
- veridicality thesis 11, 33–34, 40n7
- von Neumann, John 49, 70
- Weaver, W. 26
- Wheeler, John Archibald 42
- Wiener, N. 36, 42
- Wittgenstein, Ludwig 82–84
- Wright, Larry 61