34th ALTENBERG WORKSHOP IN THEORETICAL BIOLOGY

Causal Foundations of Biological Information

organized by
Karola Stotz and Paul E. Griffiths

July 6–9, 2017

KLI, Klosterneuburg
Austria

Welcome

to the 34th Altenberg Workshop in Theoretical Biology. The Altenberg Workshops are interdisciplinary meetings organized by the KLI in Klosterneuburg, Austria. The workshop themes are selected for their potential impact on the advancement of biological theory. Leading experts in their fields are asked to invite a group of internationally recognized scientists for three days of open discussion in a relaxed atmosphere. By this procedure the KLI intends to generate new conceptual advances and research initiatives in the biosciences. We are delighted that you are able to participate in this workshop, and we wish you a productive and enjoyable stay.

Gerd B. Müller
President

The topic

This workshop comes at the end of the three-year research project “Causal Foundations of Biological Information,” funded by the Templeton World Charity Foundation.

The source of order in living systems has been the key question at the boundary of biology and philosophy since the eighteenth century. Today, it is widely believed that living systems differ from non-living because they are driven by information, much of which has accumulated during evolution, and much of which is genetically transmitted. But there is at present no specifically biological measure of information that can underpin this vision. This project aimed to fill that gap by grounding the idea of biological information in contemporary philosophical work on the nature of causation.

The project set out to develop a measure of biological information inspired by the early theoretical insights of the co-discoverer of the structure of DNA, Francis Crick, but general enough to capture information-processing in gene regulatory networks, epigenetic information, and the emergence of new information in self-organising processes.

The workshop represents a central first step in evaluating how well these objectives have been met by presenting the main results produced by the project investigators to a selected group of philosophers of biology and theoretical biologists.

The research was conducted by: Paul Griffiths (University of Sydney); Karola Stotz (Macquarie University), Arnaud Pocheville (University of Sydney), and Brett Calcott (University of Sydney).

Aims

The project aimed to produce results that will have sufficient biological substance to be taken up by working biologists, and sufficient generality and conceptual coherence to become a standard tool in theoretical biology and philosophy of the life sciences. Theoretical biologists and philosophers of biology are the immediate target audience for the work.

Format

There will be 12 presentations, with 50 minutes allotted for each—roughly 25-30 minutes for each talk, followed by 20-25 minutes for questions and discussion. On Friday, we will kick off with an introductory statement, addressing the aims and framework of the workshop, by the organizers; on Sunday, we will end with a general discussion and publication plans. The workshop is structured into 6 sessions (‘Setting the Stage’; ‘Information & Culture’; ‘Going Formal’; ‘Networks, Signaling, & Regulation’; ‘Back to Biology’; and ‘Wrapping up/Outlook’), 4 of which will start with a presentation by one of the four team members.

To support discussion during the sessions, we encourage all participants to send a rough draft of their presentation and/or some materials that are relevant to their topic to the organizers in advance of the workshop, to be circulated among the participants.

Manuscript preparation and publication

The Altenberg Workshops in Theoretical Biology are sponsored by the KLI, the Templeton World Charity Foundation and a Macquarie University Start-up grant to Karola Stotz. In turn, the Institute, TWCF and MQ require all participants to contribute a paper to a volume edited by the organizers. Altenberg Workshop results are usually published in the Vienna Series in Theoretical Biology (MIT Press). The volume will further develop the novel ideas and concepts generated as a result of the workshop. The contributors are not necessarily limited to the original participants; they may be complemented by experts on those topics that emerge as important and may include co-authors invited at the discretion of the participants. Because of the explicit interdisciplinary nature of the effort, the outcome should be attractive to a wide range of experts in the natural and social sciences as well as in the humanities.

We expect that participants will revise their drafts as a result of our discussions at the workshop and the ensuing review process (probably “round-robin,” during which commentaries will be elicited for each paper from two selected members of the workshop). We are aiming for a November 15, 2017, date for receipt of finished manuscripts for publication. The length of the contributions should be approximately 8,000 words. The use of figures and photographs is highly encouraged. All contributions will be edited for style and content, and the figures, tables, and the like will be drafted in a common format. The editors will send specific instructions after the workshop.

Karola Stotz and Paul Griffiths

Participants

ALYSSA ADAMS	**Alyssa.M.Adams@asu.edu**

Beyond Center for Fundamental Concepts in Science
Arizona State University
P.O. Box 870506
Tempe, AZ 85287–0506
USA

BRETT CALCOTT	**brett.calcott@gmail.com**

34a Owairaka Rd, Mt Albert
Auckland 1025
New Zealand

BERNAT COROMINAS-MURTRA	**bernat.corominas-murtra@meduniwien.ac.at**

Section for Science of Complex Systems
Medical University of Vienna
Spitalgasse 23
1090 Vienna
Austria

WALTER FONTANA walter@hms.harvard.edu

Department of Systems Biology
Harvard University / Medical School
200 Longwood Avenue
Warren Alpert Building 513
Boston, MA 02115
USA

PAUL E. GRIFFITHS Paul.Griffiths@sydney.edu.au

Department of Philosophy and Charles Perkins Centre
University of Sydney
NSW 2009
Australia

RUSSELL GRAY gray@shh.mpg.de

Max Planck Institute for the Science of Human History
Kahlaische Strasse 10
D-07745 Jena
Germany

EVA JABLONKA **jablonka@post.tau.ac.il**

The Cohn Institute for the History and Philosophy of Science and Ideas
Tel Aviv University
Ramat Aviv, Tel Aviv 6997801
Israel

JOHANNES JAEGER **yoginho@gmail.com**

Complexity Science Hub Vienna
Josefstädter Straße 39
1080 Vienna
Austria

ARNAUD POCHEVILLE **Arnaud Pocheville@sydney.edu.au**

Department of Philosophy and Charles Perkins Centre
University of Sydney
NSW 2006
Australia

SONJA J. PROHASKA **sonja@bioinf.uni-leipzig.de**

Institute of Computer Science
University of Leipzig
Room 315 Härtelstr. 16-18
D-04107 Leipzig
Germany

KAROLA STOTZ **Karola.stotz@mq.edu.au**

Department of Philosophy
Building W6A, Level 7
Macquarie University
NSW 2109
Australia

EÖRS SZATHMÁRY **szathmary.eors@gmail.com**

Evolutionary Systems Research Group
MTA Ecological Research Centre
Biological Institute
Eötvös Loránd University
Budapest, H1117
Hungary

Title:

Causal Foundations of Biological Information

Thursday 6 July	Evening		
6.00 pm		Welcome reception and dinner at the KLI	

Friday 7 July	Morning	Setting the Stage	Chair: K. Stotz
9.30 am – 9.40 am	Gerd Müller	Welcome address	
9.40 am – 10:00 am	Paul Griffiths	Outline of the subject and aims of the meeting	
10.00 am – 10.50 am	Paul Griffiths	“Information in Biological Theory and Philosophy of Biology”	
10.50 am – 11:20 am	Coffee		
11.20 am – 12:10 pm	Karola Stotz	“When Is a Biological Cause a Source of Information?”	
12.10 am – 2.00 pm	Lunch	at the KLI	

Friday 7 July	Afternoon I	Information and Culture	Chair: B. Calcott
2.00 pm – 2.50 pm	Eva Jablonka	“Functional Information and Sample Selection in Learning”	
2.50 pm – 3.40 pm	Russell Gray	“Macro Matters: Cultural Macroevolution and the Prospects for an Evolutionary Science of Human History”	
3.40 pm – 4:10 pm	Coffee		
Friday 7 July	**Afternoon II**	**Going Formal**	**Chair: P. Griffiths**
4.10 pm – 5.00 pm	Arnaud Pocheville	“Crick Information: Giving Substance to Biological Information”	
5:00 pm – 5.50 pm	Bernat Corominas-Murtra	Information in Evolving Systems: From Shannon Framework and Beyond	
6.30 pm		Departure for dinner to a restaurant at the Danube	

Saturday 8 July	Morning	Networks, Signaling, & Regulation	Chair: K. Stotz
9.30 am – 10.20 am	Brett Calcott	“Information and Evolvability in Gene Regulatory Networks”	
10.20 am – 11.10 am	Walter Fontana	“Causality in (Models of) Molecular Signaling”	
11.10 am – 11.40 pm	Coffee		
11.40 am – 12.30 pm	Sonja Prohaska	“Computational Aspects of Epigenetic Gene Regulation”	
12.30 pm – 2.30 pm	Lunch	at the KLI	

Saturday 8 July	Afternoon	Back to Biology	Chair: Pocheville
2.30 pm – 3.20 pm	Johannes Jaeger	“Positional Information: On the Uses and Abuses of the Term in Developmental Biology”	
3.20 pm – 4.10 pm	Eörs Szathmáry	“Efficient Causes, Organization, Gratuity, and Biosemiotics”	
4.10 pm – 4.40 pm	Coffee		
4.40 pm – 5.30 pm	Alyssa Adams	“Mechanisms for Open-Ended Evolution in Dynamical Systems”	
6.00 pm		Free evening for exploring Vienna	

Sunday 9 July	Morning	Outlook	Chair: Calcott
9.30 am – 9.50 am	Speaker to be announced	Reflections on the Meeting I	
9:50 am – 10:20 am	Speaker to be announced	Reflections on the Meeting II	
10.20 am – 11.10 am		General discussion	
11.10 am – 11.40 am	Coffee		
11.40 am – 12.30 pm		Publication plans	
12.30 pm – 2.15 pm	Lunch	at the KLI	
2.30 pm		Departure for a vineyard tour & Heurigen	

Abstracts

Paul E. GRIFFITHS

University of Sydney

Information in Biological Theory and Philosophy of Biology

The status of information in contemporary biology is paradoxical. Many biologists are convinced that biological systems are information-based in a way that other complex physical systems are not: the development of evolved characteristics is an expression of information accumulated during evolution and inherited on via the genome. This picture of biology was influentially expressed by leading biologists like Ernst Mayr and François Jacob, both of whom regarded it as the key insight of mid-20th century biology. Today, almost every popular science presentation of biology relies on this picture. But look more closely at scientific practice and the picture lacks substance. Biology today is certainly an information science, both because it is a science of big data and because many specific models are inspired by the informational sciences, but these models do not add up to an integrated theory of biological information. Perhaps `biology is an information science' only in the sense that it uses many models that start with analogies to some aspect of communication or computing, and makes many direct applications of formalisms from the information sciences. Each of these models or applications stands or falls on its own scientific merits. They do not link together to form a single theory of biological information or a theory of life as an informational phenomenon. Leading figures in the philosophy of biology, such as Sahotra Sarkar, Peter Godfrey-Smith and one of the authors, Griffiths have endorsed this sceptical, deflationary view of biological information.

Despite such criticism, leading biologists remain profoundly attracted to this view of living systems. The late great John Maynard Smith devoted significant effort to defending it, and in his work with Eörs Szathmary argued that the ‘major

transitions' were fundamentally transitions in how information is transmitted. Similarly, the influential evolutionary thinker George C. Williams used his last book to call for a new biology focused on the `codical domain' of information rather than material entities. But this is not the biology we currently have.

Numerous biologists and philosophers have proposed accounts of information intended to vindicate the picture of biology as a science of biological information. The best known of these approaches are probably `biosemiotics' and `teleosemantics', the approach favoured by Maynard-Smith and currently ably defended by Nicholas Shea. Both take information to connote features like meaning, representation and semiosis. They seek to identify these properties, traditionally associated with thought and language, in the foundations of living systems. But this moves the concept of biological information even further away from the actual use of informational frameworks in biology. In those practical contexts the central features of information to be those captured by the standard formalisms of information theory: correlated variation and mappings between structures. In my view this 'top down analogy' approach to clarifying the concept of information in biology is fundamentally flawed. Instead, we need a bottom-up strategy starting with demonstrably useful applications of information theory and related formalisms to fundamental aspects of the operation of living systems.

It is this strategy that Stotz, Pocheville, Calcott, and I have been pursuing for the past three years. In this presentation I will give a brief, non-technical outline of how we have used causal information theory to express idea of biological specificity, and to reconstruct the thinking behind Francis Crick's 'sequence hypothesis' and 'central dogma'. I will concentrate on the philosophical motivation for using these particular formal tools, and on the strengths and limitations of approaching biological information in this way.

One of my personal motivations for pursuing this research has been to state more clearly the idea that genetic and non-genetic causes in development and sometimes informationally 'on a par'. This was key aspect of the 'parity thesis formulated by Russell Gray and myself in 1994. I will show how, if information is understood in the way our team has suggested, epigenetic and exogenetic causes can express biological specificity measurable in the same, informational currency as the specificity of genetic causes.

Karola STOTZ

Macquarie Unversity

When is a Biological Cause a Source of Information

We will discuss the usefulness of the concepts of specificity and arbitrariness to understand and justify information talk in biology. We follow Woodward (2010, 314) that “biologists tend to think of structures as carrying information when they are involved in causally specific relationships.” However, we think that the relationship between specificity and information needs documenting and explaining. We will do this by analyzing key examples of biological discourse distinguishing information from non-information by using the concept of specificity. We will elucidate the epistemic payoff for Crick introducing the idea of information to explain the process of protein synthesis with his Central Dogma and Sequence Hypothesis. We go on to explore other examples that investigate biological core processes such as the control of gene expression, cell signaling and embryonic induction, the last involving the distinction between ‘instructive’ and ‘permissive’ causes. By using our information-theoretic measure of specificity we attempt to show how the distinction between informational and other causal processes does real explanatory work. The result of this work will be an explication of the idea of information, which maintains intellectual continuity with a vernacular concept whilst making that concept more precise.

Eva JABLONKA

Tel-Aviv University

Functional Information and Sample Selection in Learning

(Nir Fresco, Simona Ginsburg and Eva Jablonka)

The paper presents a selection-based view of functional information. We claim that learned, functional information is produced through exploration processes and differential, selective, stabilization in a receiver. Following Price, we distinguish between Darwinian selection (selection among multiplying replicating entities) and Sample selection (selection without replication and multiplication of a subset from a set according to some value criterion), and argue that one can relate functional information to both types of selection processes and describe the dynamics of learning by using the Price Equation.

Russell GRAY

Max Planck Institute for the Science of Human History, Jena

Macro Matters: Cultural Macroevolution and the Prospects for an Evolutionary Science of Human History

Biological thinking and methods can be applied to both cultural micro and macroevolution. However, much of the current literature focuses on cultural microevolution. In this talk I will argue that the growing availability of large cross-cultural data sets facilitates the use of computational methods derived from evolutionary biology to answer broad scale questions about the major transitions in human social organization. I will illustrate this argument with examples drawn from my recent work on the roles of Big Gods and ritual human sacrifice in the evolution of large, stratified societies. These analyses show that although the presence of Big Gods is correlated with the evolution of political complexity, in Austronesian cultures at least, they did not play a causal role in ratcheting up political complexity. In contrast, ritual human sacrifice does play a causal role in promoting and sustaining the evolution of stratified societies by maintaining and legitimizing the power of elites. I will briefly discuss some common objections to the application of phylogenetic modeling to cultural evolution, and argue that the use of these methods does not require a commitment to either gene-like cultural inheritance or to the view that cultures are like vertebrate species.

Arnaud POCHEVILLE

Charles Perkins Centre, Sydney

Crick Information: Giving Substance to Biological Information

A theory of biological information is outlined in which information plays a substantial, causal role in living systems. There are two aspects to this theory: information as determining a choice between a set of alternatives, and information as determining the construction of a single object. The first aspect has been developed in earlier work to yield a quantitative measure of biological information that can be used to analyse biological networks. In the present paper we explore the prospects for a measure based on the second aspect, and suggest some applications for such a measure.

Bernat COROMINAS-MURTRA

Medical University of Vienna

Information in Evolving Systems: From Shannon Framework and Beyond

In this talk I will revise several of the last achievements of information theory in the study evolving systems. In its minimal abstract setting, the dynamics of such systems i) predefine constraints acting before the selective pressures, ii) is path dependent, and with not stable phase spaces, and iii) use the information in a meaningful way. According to this backbone of starting points, I will show how standard information theory can provide tools for the study of evolving – eventually open ended– systems and can explain the source of some pre-selective constraints. Further I will revise the problem of dealing with information in systems whose space of possibilities evolve and whose microscopic dynamics is path dependent. As we shall see, the unicity of the fundamental functional (the entropy or the information divergence) is broken and a zoo of different functionals emerges, depending on which assumptions are made over the underlying dynamics. Finally, I will present a minimal communication system of autonomous agents which generalises the standard schema of information theory and accounts for a minimal form of meaning transmission. All these results show that there is room for the construction of a much more general theory of information encompassing the abstract properties of evolving complex systems. Nevertheless, a final word of caution is needed: The need for a formal framework generalising information theory to study such systems must match to a real demand, and it must contribute critically to some prediction or demonstrate its unavoidability in some fundamental piece of the general theory. If it is so or not, it is still an open question.

Brett CALCOTT

University of Sydney

Information and Evolvability in Gene Regulatory Networks

A central claim in evo-devo is that gene regulatory evolution plays a distinctive role in the evolution of novelty, for it allows the existing functionality to be coopted and recombined in new ways. Using a model of gene regulatory evolution, I explore two ways that information theoretic measures of causation can clarify such claims. First, I use an information measure to identify functional (rather than topological) modules in gene regulatory networks. Second, I show how an information measure can capture an intuitive aspect of evolvability, measuring how readily these same modules can be coopted to provide novel functionality.

Walter FONTANA

Harvard University

Causality in (Models of) Molecular Signaling

Over the last decade, we developed a rule-based framework for modeling complex systems of protein-protein interaction as they occur in signaling and assembly. The approach is fashioned after chemistry, but defined at a level of abstraction more suited for molecular biology. Since rules only express local mechanisms of interaction, the question arises: What, exactly, is a pathway? In an attempt at answering this question, I will discuss notions of influence and causality that are informed by concurrency, where the latter refers to situations in which the ordering of certain events does not matter.

Sonja J. PROHASKA

Leipzig University

Computational Aspects of Epigenetic Gene Regulation

Genes and environmental factors have been conceived as the determinants of phenotypic characteristics of complex multicellular organisms. However, how the genotype is translated into the phenotype is still poorly understood at a molecular level.

Over the last years, *epigenetics*, a layer of chemical modifications sitting "on top of" the DNA regulating when and how genes are expressed, was put forward. Meanwhile, these processes have been accused of serving functions ranging from genome indexing to creating an internal representation of the outside world.

To get an understanding of how a regulatory system, though theoretically of extraordinary power, could serve all these purposes, we address the questions: how did it come about, and what does it do?

In this attempt, drawing parallels to concepts from computer science will provide useful.

Johannes JAEGER

Complexity Science Hub, Vienna

Positional Information: On the Uses and Abuses of the Term in Developmental Biology

Ever since Lewis Wolpert introduced the term 'positional information' in 1968, developmental biologists have been using it in a great number of contexts and with a great range of meanings which often remain ill-defined. In my contribution, I will trace some of its uses in the literature and highlight a number of obvious problems with the concept. I will also critically review more recent work, which is explicitly based on an information-theoretic approach. This review of the concept and its history illustrates many ways in which the term can be used productively, but also many of the pitfalls that await those who do not reflect on its meaning carefully enough.

Eörs SZATHMÁRY
Eötvös Loránd University, Budapest & MTA Ecological Research Centre, Tihany

Efficient Causes, Organization, Gratuity, and Biosemiotics

Information in biology has a different role than, say, in chemistry. In the physical sciences many would subscribe to the view that information is "anything that can be known" (by whom?). In biology, we tie the useful notion of information to agency. Of course, during the origin of life, some information as defined by chemists has become also information as understood in biology, due to the appearance of replicators and reproducers. At the cellular level, there is considerable material overlap between parent and offspring. There are other things than classical genetic information that is being transmitted. What is it? A kind of minimal organization. This minimal organization is a set of autocatalytic entities that are informationally independent of genes to various degrees during evolution. This set includes some autocatalytic metabolic cycles and genetic membranes. As Gánti recognized, cellular organisation is an autocatalytic supercycle composed of qualitatively different component autocatalytic entities. Rosen was first to propose that such organizations have closure in terms of efficient causes. Such a closure contains irreducible and indispensable information.

Given an organization, signals transmitted between agents can be iconic, indexical and symbolic. The latter is related to the concept of "gratuity" in molecular biology. The genetic code and natural language can be arguably regarded as symbolic systems, but they are unlikely to have started so. I shall also examine the question of "biological codes" as advocated by Barbieri in this regard.

Alyssa ADAMS

Arizona State University

Mechanisms for Open-Ended Evolution in Dynamical Systems

Open-ended evolution is a defining feature of biological systems, but has been challenging to reproduce *in silico*. We recast the problem in dynamical systems theory, providing simple criteria for open-ended evolution based on two hallmark features: unbounded evolution and innovation. As a case study, we implement novel variants of cellular automata (CA) where the update rules are allowed to vary with time in three alternative ways. We find that state-dependent dynamics statistically out-performs other mechanisms, and is the only mechanism to produce open-ended evolution in a scalable manner, essential to the notion of ongoing evolution. In addition, CA that only utilize reversible rules are capable of producing far more open-ended dynamics, suggesting information-preservation in states is a crucial feature for open-endedness. This analysis suggests a new framework for unifying mechanisms for generating OEE with features distinctive to life and its artifacts, with broad applicability to biological and artificial systems.