Biol Philos (2011) 26:195–200
DOI 10.1007/s10539-011-9257-3

DISCUSSION NOTE

Response to commentaries on "The Transmission Sense of Information"

Carl T. Bergstrom · Martin Rosvall

Received: 1 January 2011 / Accepted: 14 February 2011 / Published online: 29 March 2011

First of all, we wish to thank the respondents for their comments, both formal and informal, on our treatment of the transmission sense of information (Bergstrom and Rosvall 2010; Godfrey-Smith 2010; Maclaurin 2010; Shea 2010). Through this type of dialogue, both as we wrote the original paper and during the present process of commentary, we have been able to substantially improve and refine our approach to the problems of information in biology. For this we are very grateful.

Second, we want to stress that when we talk about information in biology, our use of the term "information" is not intended to be metaphorical. We mean it literally. We are not saying that it is *as if* genes carry information; we are saying that genes *do* carry information, and that their role as information carriers is clear from their structure and function.

Non-semantic information?

Each of the commentators recognizes the transmission sense of information as a legitimate non-semantic account of biological information. One concern that each raises, and Shea in particular highlights, is where this leaves the notion of semantic

C. T. Bergstrom (✉)
Department of Biology, University of Washington, Seattle, WA 98195-1800, USA
e-mail: cbergst@u.washington.edu
URL: http://octavia.zoology.washington.edu/

C. T. Bergstrom
Santa Fe Institute, 1399 Hyde Park Rd., Santa Fe, NM 87501, USA

M. Rosvall
Integrated Science Lab, Department of Physics, Umeå University, 901 87 Umeå, Sweden
e-mail: martin.rosvall@physics.umu.se

content. Does the transmission view imply that information is non-semantic? Are we denying that biological information has semantic content? No, and no.

In our view, the transmission sense of information does not refer to a *kind* of information, distinct from semantic information and devoid of semantic content. Rather, the transmission view is *justification* for identifying something as information, distinct from semantic justifications. In our view, it is entirely possible that informational objects in the transmission view will be imbued with semantic content. In fact we would be surprised if the vast majority of evolved informational structures did not carry some sort of semantic content. Why would the information transmission have evolved if that information did not "mean" something to some receiver?

Thus we readily accept that DNA sequence carries semantic information about primary amino acid sequence at the very least, and perhaps about higher-level aspects of phenotype as well. Our point is that we do not have to resolve those disputes about the reach of semantic representation in order to be certain that DNA is an informational molecule. Following Claude Shannon (Shannon 1948), we see a decided utility to thinking about information while remaining agnostic to its semantic content.

But this is a matter of some concern to our discussants. Even if we do not deny that biological information has semantic content, they rightly observe that our approach deliberately ignores that semantic content. Because to many biologists and philosophers the semantic content is the "interesting part," these authors worry that the transmission view of information is one that ignores the most interesting aspects of biological information. For example, Godfrey-Smith expresses surprise that we say so little about what the semantic content of genes might be. There is a reason for this. The entire point of the transmission view is that messages can be identified as messages because of how they are structured and how they function with respect to transmission. We do not need to appeal to an understanding of what they say and how they function semantically.

This is not to say that we dissent from the majority view that semantic meaning is interesting. Moreover we agree that the question of genes' semantic content cannot be set aside by science if we want to understand how biological inheritance works. We believe we have shown that an adequate naturalistic account of information content in the genome need not be semantic, but at the same time we freely admit that further work remains to develop an adequate semantic account. We see how this concession might feel unsatisfactory to those more interested in semantic theories than in information for its own sake, but at the same time we feel that providing any sort of justifiable naturalistic account of biological information represents a significant advance.

Maclaurin's library

In his commentary, Maclaurin makes an analogy between our analysis of information in biology and a library. He worries by pursuing the transmission view, we have committed to viewing the library from a cataloguer's perspective—

an art "that only a librarian could love"—whereas most sensible people are bored by bibliographic classification and care about libraries because they are interested in the actual contents of books.

But take the perspective of the proverbial anthropologist from Mars. If the anthropologist wants to understand how humans keep track of so many ideas to be retrieved and recombined and reused, the anthropologist will need to study the human institution of libraries. But how will the Martian anthropologist identify a library? How will she avoid limiting herself to wood-pulp technologies, and manage to see that the World Wide Web functions very much as a library of a sort? How will she manage to distinguish between the New York Public Library and a warehouse full of thousands of unsold copies of dime-store novels, waiting to be recycled into toilet paper?

The answer is that a Martian anthropologist can observe the processes by which items are indexed, catalogued, stored, and transmitted, and received, and retrieved. We freely admit that these processes are not the most *interesting* thing to be found within the library—the anthropologist won't learn all that much about humanity by studying the Dewey Decimal system—but these processes are highly *diagnostic of* a library. Our anthropologist shouldn't try to identify what serves as a library and what does not by walking into a room, picking up a book, and starting to read—even though that might be exactly what she would want to do to learn about humanity once a library has been found.

This confusion between what is interesting and what is diagnostic appears to underly Maclaurin's fundamental concern: "In short then, Bergstrom and Rosvall's account of biological information is available and coherent. It will be of great interest to those whose main interest is gene expression. However, it remains to be seen how relevant it will turn out to be in the life sciences as a whole and in wider academic debate."

The transmission sense of information does not aim to enumerate the many wondrous uses of information in living systems or the processes by which it accumulates in genomes over evolutionary time. It simply offers the tools to diagnose biological information where ever it can be found, and dispels arguments that no solid theoretical grounding can be given to the use of the term in evolutionary biology.

In addition, it strikes us that Maclaurin underestimates the importance of storage and cataloging technology when he asserts that "Systems used by library cataloguers have no effect on the semantic information contained in books." Perhaps his intent is merely to state the obvious fact that once a book is written, its meaning tends to be changed minimally, if at all, by the cataloging process. This is fair enough. But in relating the library metaphor to biological information, it is key to note that the storage and cataloging technologies have a vital role in determining what kinds of semantic information a system can afford to store and access. One can imagine that the transitions from clay to paper and from handwriting to printing press caused quantum leaps in the production efficiency of text and in the storage capacity of libraries, driving concomitant advances in cataloging and information retrieval. As a result, new information niches opened up: libraries could afford the space and publishers the effort to record a broader range of information, beyond a narrow collection of classics and the ever-present Bible. The transition from paper to magnetic storage has had similarly profound consequences. Again we have seen a quantum leap in storage capacity. But more importantly, the hyperlinks among documents stored on a vast computer network—originally designed to solve problems of document delivery—have turned out to provide an even more valuable service in the domain of document discovery via Google's Page Rank algorithm and other network data mining approaches. Even the most obscure sources of information about the most abstruse subjects now merit inclusion in the *World Wide Library*, because advances in indexing technology allow them to be found.

We venture that something similar has happened in the history of life. Many of Maynard Smith and Szathmary's major transitions in evolution—the transition from RNA to DNA, the evolution of epigenetic markup, or the evolution of cultural transmission—are precisely the sorts of advances in storage and "cataloging" technology that Maclaurin denigrates (Maynard Smith and Szathmáry 1995). Each has opened up new niches for life on the planet, making it possible to store and to process not only greater amounts of information, but also information *about* things that were not previous feasible to address.

A small but important refinement

The present dialogue led us to an important refinement of our original definition of transmission-sense information. We had originally situated information in the "sequence properties" of objects. In a very brief aside from his main argument, Godfrey-Smith questions whether it is always the sequence properties that convey information. He is right to object. We should not make too much of the analogy between Shannon's serial communication channel and the sequential serial structures of nucleotide bases in DNA and of amino acids in proteins. Information transmission need not be sequential. For example, we could communicate perfectly well by exchanging unordered hands of five playing cards from a deck of 52. To encompass this sort of scenario while including sequence properties as a special case, we suggest that "sequence properties" should be expanded to include all types of "combinatorial properties", ordered or unordered:

> **Transmission view of information:**
> An object X conveys information if the function of X is to reduce, by virtue of its combinatorial properties, uncertainty on the part of an agent who observes X.

Messages and memories, germlines and soma

In his *Mathematical Theory of Communication*, Shannon illustrated the mathematics of information with a basic model of communication as through a simple telegraph relay (Shannon 1948). In this telegraph schema, a sender encodes a message, and transmits it through a channel to a receiver who decodes and interprets it.

A central concern in Godfrey-Smith's commentary is the following: Is there a clean mapping from informational processes in biology onto the telegraph schema? Or do we even need one? This is an important issue and one that our account, or any other account grounded in Shannon's information theory, will need to address.

For many cases, the mapping may be relatively straightforward. Indeed, a more-or-less direct translation of Shannon's schema may carry us a long way toward understanding the *origins* of biological information. But some of the most interesting aspects of information use in modern organisms seem to defy simple parallels to Shannon's basic setup. Fortunately, the conceptual heart of information theory lies not in the telegraph schema, but rather in the fundamental mathematics that describe the mechanics and limits of information transfer. We treat these two issues in turn.

First, let's consider the reach of Shannon's basic schema. As Godfrey-Smith notes, memory can be viewed as involving a transfer of information, where the sender and receiver are the same individual considered at different points in time. This observation appears sufficient grounding for the use of something very much like the telegraph schema in conceptualizing the origins of biological information.

The reason is this: The last universal common ancestor (LUCA) of all living things was a unicellular life form that arose approximately 3–3.5 billion years ago. It was not until 1.2 billion years ago that the first simple multicellular lineages arose, and it was a mere 580 million years ago that the Avalon explosion provided this planet with its first large and complex multicellular organisms (Shen et al. 2008). Because all living things share a common genetic code, it is clear that DNA-based hereditary information arose prior to LUCA. Indeed, DNA served to transmit information from unicellular generation to unicellular generation for more than a billion years prior to the evolution of multicellular differentiation. Thus the origin of DNA as an informational molecular clearly predates the origin of the developmental-regulatory functions that it currently fulfills in metazoans, and about which Godfrey-Smith and others have expressed so much concern. For early unicellular lineages, the "memory channel" schema may map very closely to the biological role of DNA. With the transmission sense of information, we can explain information flow via DNA along the germ line, without resorting either to semantic theories of meaning or to mere correlational interpretations of information.

Second, let's move past the basic Shannon schema to look at how information flows not only between generations along the germ line, but also within a single generation through the process of multicellular development. How should we conceive of this type of system? Is the sender an individual organism and the receiver one of its offspring, or is the sender a single cell within an individual and the receiver a subsequent member of its cell lineage? We certainly need to deal with this issue of how to interpret notions of senders and receivers, encoders and decoders, in multicellular organisms. This has been where the action has been in previous debates on the subject, and there is more work to be done in this domain. The principal problem is that under our definition of information, biological information seems to be flowing in two orthogonal directions: "down" through the generations, and "across" through lineages of semantic cells. Nothing like this appears in the basic telegraph schema that we tend to think of when we think about Shannon's theory.

Fortunately, Shannon's theory goes far beyond the telegraph model of serial information transfer through a single unidirectional channel. Much of the excitement in contemporary information theory research derives from extending Shannon's approach in such directions (Cover and Thomas 2006). Feedback channels allow two-way communication between signaler and receiver and afford various schemes for validating that messages have been successfully received. Multiple-access channels are shared by numerous signalers who may have no way to coordinate their actions and whose messages may therefore interfere with one another. Network information theory—a critical area in designing cellular communication networks, among other applications—addresses how information can move through entire networks of signalers and receivers, and deals with issues such as how interference can be avoided and how intermediaries can effectively relay messages to downstream receivers.

The key point is that the telegraph schema is just a diagram that Shannon chose for illustrating simple cases of his mathematical theory. It is in the underlying mathematical framework, not in the simple telegraph schema, that we are situating the stuff that we call information. The reason that information theory has been an enormous success in the engineering world is that its mathematical truth affords endless applications beyond those that Claude Shannon himself may have envisioned. When structures become more complicated than can be represented by the basic telegraph schema, we can keep using the theory. Just as the theory has extended beautifully to more complicated scenarios in telecommunications engineering and computer science, we expect the same for application to biological heredity and development. We believe that the transmission sense of information has the capacity to lead us in interesting and productive directions of inquiry into the origin and nature of biological information, and we look forward to pursuing these directions in the future.

Acknowledgments CTB was supported in part by NSF ATB-1038590.

References

Bergstrom C, Rosvall M (2011) The transmission sense of information. Biol Philos. doi:10.1007/s10539-009-9180-z

Cover TM, Thomas JA (2006) Elements of information theory, 2nd edn. Wiley, New York

Godfrey-Smith P (2011) Senders, receivers, and genetic information: comments on Bergstrom and Rosvall. Biol Philos. doi:10.1007/s10539-010-9206-6

Maclaurin J (2011) Commentary on "The transmission sense of information" by Carl T. Bergstrom and Martin Rosvall. Biol Philos. doi:10.1007/s10539-010-9233-3

Maynard Smith J, Szathmáry E (1995) The major transitions in evolution. Oxford University Press, Oxford

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27, 379–423, 623–656

Shea N (2011) What's transmitted? Inherited information. Biol Philos. doi:10.1007/s10539-010-9232-4

Shen B, Dong L, Xiao S, Kowalewski M (2008) The Avalon explosion: evolution of ediacara morphospace. Science 319(5859):81–84