Biol Philos (2011) 26:177–181
DOI 10.1007/s10539-010-9206-6

DISCUSSION NOTE

Senders, receivers, and genetic information: comments on Bergstrom and Rosvall

Peter Godfrey-Smith

Received: 18 March 2010 / Accepted: 25 March 2010 / Published online: 22 April 2010
© Springer Science+Business Media B.V. 2010

It is tempting to say that genes contain messages or instructions of some kind. For many, "tempting" is not the right word; this is something that modern biology has discovered and described. But the boundary between literal and metaphorical description is hard to work out in cases like this. Some of those who would insist on the enormous value of symbolic and computational description of genetic systems may find themselves adding that the goodness here is ultimately a goodness of metaphor. As I understand them, Bergstrom and Rosvall ("B&R") are quite literal-minded in their approach. They see their information-theoretic analysis as describing real features of genetic systems. They argue that the previous literature on this topic has failed to see these features because it has not approached the question in the right way. The right approach is decision-theoretic: "information theory is a decision theory of how to package information for transport, efficiently" (Bergstrom and Rosvall this issue).

Within that framework, B&R assert the following view:

Transmission view of information:

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

B&R can be seen as working within a more general "sender-receiver" approach to questions about information, meaning, and related topics. If the object X in their formulation above has the function to convey information, then not just its reception but also its production must have occurred in accordance with some kind of design (either deliberate or evolutionary). So X is being used to achieve coordination between one "agent" and another. I will accept this approach here, and will also set aside some minor issues (e.g., is it always the "sequence" properties of X that

P. Godfrey-Smith (✉)
Harvard University, Cambridge, USA
e-mail: pgs@fas.harvard.edu

function to reduce uncertainty?) The comments I will make are arrived at within B&R's own framework. I will emphasize two points. First, on a sender-receiver view like this, a message is only as real as its receiver or reader—the "agent who observes" in B&R's formulation. This is not to say that the physical vehicle of the message has some strange dependence on the existence of a receiver, but for a physical object to count as a message, there must be a receiver present at least in normal circumstances. Second—or perhaps equivalently—the particular *content* of a message depends on the nature of the sender and receiver. A message has the function of reducing uncertainty in a receiver of a certain kind. There is no point in positing a message that the receiver in question cannot use, apply, or comprehend. Talk of messages and content is talk of sender-receiver coordination.

In the case of genes, who is the sender and who is the receiver? B&R say: "In biology, transmission genetics is the study of inheritance systems, not the study of transcription and translation, and genetic transmission is the passing of genes from one generation to another". So the sender in each case is an organism, and the receiver is another organism in the next generation. What sort of content do the messages have? "[G]enes are transmitted from parent to offspring in order to provide the offspring with information about how to make a living (e.g., metabolize sugars, create cell walls, etc.) in the world".

Certainly this case is rather different from the most familiar sender-receiver systems. In familiar cases, the sender and receiver are pre-existing things, spatially separated, and the message is sent from one to the other. In the genetic case, the receiver does not exist when the message is "sent." Instead the receiver is brought into being by a causal process which includes, on B&R's view, the reading of the genetic message. I do not think these disanalogies are problems, however. The framework of information theory can be applied outside the most familiar cases. For example, B&R note that messages can be sent through time as well as space. This means that an information-theoretic approach can be applied to *memory*; the operation of human memory is like sending messages through time by laying down traces in the brain. These traces can be later consulted—the self at a later time is the "receiver." We can also imagine a case in which a sender sends out a machine which includes both a message on a hard drive and a device which reads the message and uses it to rebuild the entire machine.

In the biological case, we have the following events. A zygote is formed by the fusion of gametes. The cell divides. Genes are transcribed and translated. As the cells divide, different genes come to be read in different cells, producing the differentiated structures seen in a multicellular organism in the new generation. Looking for senders and receivers here, let's start at the level of individual cells. Here we find processes that do look like the "reading" of genes, in protein synthesis, and this may be seen as the use of a message. Who is the sender of these messages? We might say that the sender is always an earlier cell in the same lineage. Some might say that a cell lineage uses DNA as a form of memory—it sends messages through time. If we are talking about the zygote itself, then two cell lineages have come together to produce it. In the case of the zygote, it is certainly tempting to zoom out from the cellular level and say that the parental organisms are senders. However, I do not think it makes sense to say that the whole organism

receives the message. The "reading" of the message that has been recognized so far is a process occurring within individual cells. As a result of all the low-level reading, the cells dividing, and the differentiation of which messages are read in which cells, a new multicellular organism comes to exist and develops. But each message being read is specifying no more than the sequence of an individual protein molecule (or just an RNA). That is the information that the readers in the cells can make use of. The machinery of protein synthesis in a cell does not read a message like "Have fingers!" or "Here is how to make a kidney!"

What is the content that B&R think is being sent, according to their transmission-based view of genetic information? I gave a quote earlier: "genes are transmitted from parent to offspring in order to provide the offspring with information about how to make a living (e.g., metabolize sugars, create cell walls, etc.) in the world". It is surprising that this is nearly all B&R say on the topic. B&R insist that genes contain a message, but do not say much at all about what the message is.

B&R give reasons for saying little on this topic:

> When information theorists think about coding, they are not thinking about semantic properties. All of the semantic properties are stuffed into the codebook, the interface between source structure and channel structure, which to information theorists is as interesting as a phonebook is to sociologists. When an information theorist says "Tell me how data stream A codes for message set B," she is not asking you to read her the codebook. She is asking you to tell her about compression, channel capacity, distortion structure, redundancy, and so forth. That is, she wants to know how the structure of the code reflects the statistical properties of the data source and the channel with respect to the decision problem of effectively packaging information for transport.

I do not think this addresses the issue, however. An information theorist in their normal working context might take for granted that there are messages being sent in a particular case, and might focus on some quantitative features of the transmission system which would be present regardless of the exact contents of the messages. But a complete story about any such case would have to include specification of what the messages say. If there is a measurable reduction of uncertainty going on, then there *is* uncertainty that is being reduced—uncertainty about something beyond the signals. So the question of the content of the genetic message cannot be set aside.

Given that, let's look more closely at what B&R did say—at the idea that genes contain information about things like how to "metabolize sugars, create cell walls, etc." Above I compared two kinds of message. One kind contains information about the primary structure of proteins. The other contains information about whether fingers are worth having, or how to build a kidney. B&R's examples are "in between" these, as far as their scale in the organism is concerned. Was this deliberate? Or were these examples chosen more-or-less randomly, where B&R do think that the genetic message includes information about (say) how to build a kidney? For me, there is a big difference between the two kinds of content. It is a fact about development in organisms like us that there is no reader reading a "Make

fingers!" message. All the reading that is going on is at a lower level. Development is a process in which a lot of low-level reading, and a lot of low-level quasi-logical cascades of gene regulation, occur, and as a result of that (among other things) an organism with fingers and kidneys comes to exist. There is no Grand Central Reader of an overall developmental message, of the kind I imagined in the case of the machine sent out into space with a message on a hard drive telling it how to rebuild itself.

Some writers in this area argue that there *are* readers at a higher level in organisms, even though they are not visible as definite parts of the machinery. The example I have in mind is Shea (2007).[1] He thinks that a "developmental system" is the reader or "consumer" of the genetic message. This, for Shea, is not a situation where we have a separate "developmental system" in each cell. Rather, visible across the whole organism there is a single developmental system which reads all the DNA that was present in the zygote, and which has been copied and distributed across many cells. This system reads a message which represents whole-organism phenotypic characteristics. I think, in contrast, that this "developmental system" is not a genuine reader of a message. It is an abstraction, and a reader should be a physical device. There are reader-like mechanisms in organisms, visible within cells, but putting a lot of these cell-level readers reading cell-level messages together does not yield an organism-level reader reading a message specifying phenotypic features found at the whole-organism level.

I am curious whether B&R would agree with Shea on this point, or with me. Do B&R hold that an offspring organism, or something within it, is a reader of messages whose contents represent or specify features of the whole organism? Or are the only messages carried by the genes ones that involve within-cell processes and products, such as proteins? This is a question that should have a definite answer on the B&R view. While on this point, I note that in their (very interesting) discussion of the design of the genetic code, and the protein-sample inheritance system imagined in my 2000 paper, the points B&R make are all compatible with the view that genes contain cell-level "micro-contents" only—contents which specify the primary structure of proteins.

B&R hold that it is an objective fact that DNA has been designed by evolution to transmit information. They make a strong case for something at least *like* this view. In other new work, sophisticated versions of the idea that the information-theoretic description of genes is a useful *fiction* have been developed (Levy forthcoming). As noted earlier, this is an area where it is hard to work out where literal use of theoretical language stops and metaphor begins. Perhaps what is going on is this. We are struggling to apply a certain vocabulary and framework to a system which has *partial* analogies to the cases that are "home uses" for that vocabulary and framework. In response, people engage in a mixture of literal theoretical extension and metaphor, with unclear boundaries between the two.

[1] Possibly Sterelny et al. (1996) and Maynard Smith (2000) are in this category also, though their "teleo-semantic" proposals were not expressed with explicit use of the idea of a reader or consumer of the inherited message.

References

Bergstrom CT, Rosvall M (this issue) The transmission sense of information
Godfrey-Smith P (2000) On the theoretical role of 'Genetic Coding'. Philos Sci 67:26–44
Levy A (forthcoming) Biological information: a fictionalist account. To appear in *Noûs*
Maynard Smith J (2000) The concept of information in biology. Philos Sci 67:177–194
Shea N (2007) Representation in the genome, and other inheritance systems. Biol Philos 22:313–331
Sterelny K, Smith K, Dickison M (1996) The extended replicator. Biol Philos 11:377–403