Sender-Receiver Systems Within and Between Organisms

Peter Godfrey-Smith
City University of New York

Presented at PSA 2012, San Diego, as part of a symposium on "Signaling Within the Organism," with Nicholas Shea, Rosa Cao, Brett Calcott, and Rory Smead.

I introduce a theoretical framework, due originally to Lewis and Skyrms, used in several talks in the PSA symposium "Signaling Within the Organism." I contrast sender-receiver systems as they appear *within* and *between* organisms, and as they function in the bridging of *space* and of *time*. Memory can be seen as the sending of messages over time, communication between an earlier and later self. I then compare psychological and genetic memory with respect to their relations to the sender-receiver model, arguing that some puzzles concerning genetic information can be resolved by seeing the genome as a cell-level memory with no sender.

1. Introduction

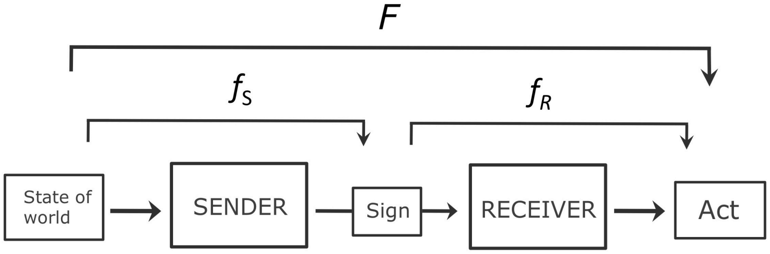
Our understanding of meaning and representation is being transformed by new work that studies signaling games of the kind described by David Lewis, but treats them from an evolutionary point of view. This approach was introduced by Brian Skyrms in *Evolution of the Social Contract* (1995) and *Signals* (2010).

The basic case that the Lewis model covers is communication between two people or other organisms. But there are ways to recognize communication and signaling within an organism as well. This PSA session is about the application of sender-receiver models to these within-organism interactions. This paper will introduce the relevant models, discuss their within-organism application in general terms, and then discuss some specific cases and problems.

2. Sender-Receiver Models

Lewis (1969) wanted to understand "conventional" signaling. He did so like this. Imagine there are two rational agents. One is a "communicator" or (as I will say) *sender* who can see the state of the world but cannot act except to produce signals of some kind. The "audience" or *receiver* can only see these signals, but can act in a way that affects them both. Lewis assumed *common interest* between the two agents: they have the same preferences for what they want done in each state of the world. Then if the sender sends distinctive signals in each state of the world, and the receiver uses these signals to prompt appropriate actions, such a "signaling system" can be a *Nash equilibrium* – neither side has any incentive to change (unilaterally) what they are doing.

More formally, the sender and receiver each follow or implement rules: f_S for the sender and f_R for the receiver. The sender's rule maps states of the world to signs; the receiver's rule maps signs to acts. Composed, the result is a function from states to acts, F, which may or may not correspond to the preferred mapping for both agents, F^*.



f_S : sender's rule, maps states of the world to signs.
f_R : receiver's rule, maps signs to acts.
F : the resulting mapping from states to acts.
F^*: the mapping from states to acts preferred by both agents.

Figure 1: Lewis Sender-Receiver Model

This situation can be described as a division of labor: the sender acts as the receiver's eyes, the receiver acts as the sender's hands. Such a division of labor requires communication between the two sides, so signs appear in the middle. This set-up can also be seen as a solution to a particular kind of problem of control: pairing behavior usefully with states of the world, dealing with the *environmental contingency of effective action*.

Lewis assumed that the sender and receiver are rational agents linked by common knowledge. In his 1996 and 2010 books Skyrms gave an evolutionary, naturalistic treatment of the Lewis model: rational choice was replaced by natural selection and also by simple forms of learning. Signaling systems can evolve and be evolutionarily stable.[1]

The Lewis model, as Skyrms emphasizes in *Signals*, has connections to information theory. When Shannon in 1948 gave a quantitative treatment of the transmission of information, he set out by imagining a *transmitter* and *receiver*, connected by a *channel* over which signals pass.

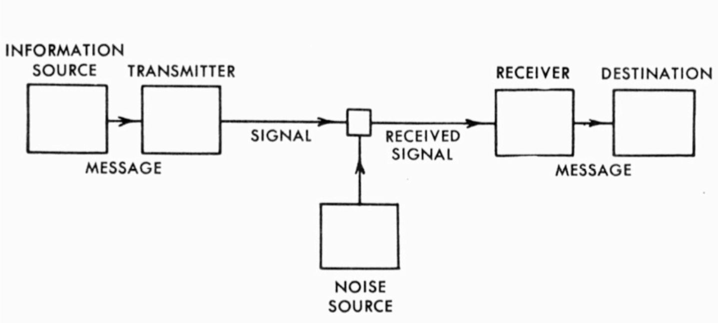


Figure 2: Shannon's 1948 diagram of a "general communication system"

He introduced measures of the amount of uncertainty associated with any variable – Shannon's *entropy* – and various measures of the degree of association between two variables – the extent to which the value of one variable predicts the value of the other. One of these measures is now called the *mutual information*.[2]

Shannon's model was set up by imagining something like a sender and a receiver, but those roles are not essential to his framework. Mutual information is ubiquitous in physical systems. The Lewis model, though, describes one way that it can come to exist: by the shaping of senders' behaviors. The sender is responsible for the informational relations between signs and states of the world, but the sender's payoffs, and hence their choices, are dependent on what the receiver does with those signals. Sender and receiver

---

[1] For an evolutionary treatment of the Lewis model see also Huttegger, Skyrms, Smead and Zollman (2010). For its philosophical significance see also Harms (2004).

[2] The *Shannon entropy* of a discrete random variable X which has possible values x_1, x_2, x_i... is $H(X) = -\sum_i P(x_i) Log_2 P(x_i)$. The *mutual information* between two variables X and Y is $H(X)+H(Y)-H(X,Y)$. Here $H(X,Y)$ is the "joint entropy" of X and Y, which can be thought of as the entropy of the distribution of combinations of X and Y values.

behaviors co-evolve, and signs have their informational properties shaped as a consequence. When the sender's rule maps states to signs one-to-one, for example, the signs carry as much information about the world as is possible. But the sender will only perform such a mapping if the receiver will do something useful – from the sender's point of view – with the information supplied. If sender and receiver want different acts performed in each state of the world, then any information about the world contained in signals will be exploited by the receiver. Then the sender should stop sending informative signals and the receiver, in turn, should stop paying attention to them.

There are cases where information in a sign-like object is useable by a receiver without a sender being involved at all. These include what biologists call "cues" – unsent or naturally occurring signs. Using clouds as a sign of the weather is an example. From the point of view of the model, these are cases where the production process behind a sign is unaffected by feedback from the sign's use and its consequences. The same might apply on the receiver side. It appears that in cases where there is no sender shaping the signs in a way affected by the receiver's actions, there is no way for cues to become complicated and elaborate. Complexity in signs cannot *accumulate* by a selection process. This does not limit the usefulness of cues, if a reader is smart enough. A smart reader can infer a vast amount from traces that are left fortuitously. (Consider Laplace's demon.) But this requires powerful inferential capacities; it cannot be done by a simple, mechanical reader device.

Sender-receiver set-ups arise in a wide range of contexts. At least three kinds of generality can be distinguished. First, the model can be applied to signaling *between* and *within* organisms. Second, different selection processes operating on various time-scales can shape sender and receiver behaviors, including evolution by natural selection, reinforcement learning, differential imitation, and deliberate choice. Third, signs can be used by senders and receivers to bridge both *space* and *time*. This gives the model a connection to *memory*, both in psychological and other forms.

There is a separate question about how broadly the Lewis model applies. As discussed above, the relations between a Lewisian sender and receiver enable the solution of a particular problem of control. Effective pairing of actions with states of the world is achieved by the sender tracking those states making information about them available to the receiver. Consider, in contrast, a "cox" in a boat calling the stroke for two rowers. The sender has no private knowledge, and the point of the calling is purely to coordinate the actions of the rowers. It does not matter when they row, as long as they row at the same time. (The cox is like a "central pattern generator" in the brain.) This *purely coordinative* signaling does not fit the Lewis model.[3] There are also mixed cases. If instead of a cox, one of the *rowers* calls the stroke, then he may have private knowledge – of the timing of his upcoming stroke – and formally the Lewis model can be applied, by treating the sender's intention to row as a "state."[4] The question arises whether this assimilation obscures the fact that the function of the call is the same in the coxless case and the case with the cox – the function is coordination of actions – and the fact that coordinative and fact-tracking signaling are distinct activities, though they may also appear in combinations.

However those issues are handled, the model brings with it a particular orientation to questions involving signs: focus on the shaping and stabilization of sender and receiver behaviors, rather than on peculiar properties of the mediating entities themselves. Here "senders and receivers" include writers and readers, producers and consumers (Millikan 1984), inscribers and interpreters, publishers and subscribers, of many kinds. As Lewis said in 1969, if you know what the sending and interpreting rules are and why they are the way they are, nothing has been left out of an understanding of the semantic properties of signs in that system. The way to understand sign-related phenomena is to investigate the co-evolution of sender and receiver behavior, the two-sided phenomenon of sign use.

This session is concerned with a particular slice of the total set of phenomena to which the framework can be applied: signaling within the organism.

3. Representation and Memory

A sender-receiver set-up of the Lewisian kind solves a particular problem of control: coordinating actions with states of the world. If one part of a system is receptive and another part can act, and there is a high degree of common interest, signaling can be used to solve the problem of determining how to act. This distinction between roles can be

---

[3] For material in this paragraph I am indebted to Kevin Zollman (in discussion).

[4] I say "may" because the caller might be such that their call prompts their own stroke, rather than reflecting the antecedent fact of their intention.

applied to the control of behavior by an organism. Both the problem and the solution can be laid out in two dimensions, in fact. First, the relevant divide is seen in the relation between an organism's senses and effectors (as discussed in Cailin O'Connor's talk at the 2012 PSA meeting).[5] Here signaling overcomes a spatial divide. Second, the division is present over time. Present selves have access to information that will be useful to the guidance of action by future selves, so present selves lay down marks that can be consulted later. Memory is sending messages over time.

The set-up characterized by the sender-receiver model appears in both clear, paradigm forms and in partial, marginal, or semi-submerged cases. In different contexts different kinds of partial cases arise. In the between-organism case, complications and non-Lewisian phenomena arise especially from the divergent interests and multifarious agendas of communicating agents. Where interests diverge, there may be a refusal to send informative signs, also cycles and chaotic dynamics (Wagner 2011).

In the within-organism case, common interest is not as much of an issue, though there are exceptions.[6] Complications arise especially from the entanglement of roles that are treated as distinct in the model. The sensorimotor axis of an organism can be seen as a realization of a sender-receiver structure in at least a rough way: the problem of coordinating acts with the world is being solved by coordinating receptive and active parts of the organism. But except in very simple and perhaps very modular cases, there is poor separation between sender and sign; there is a network of downstream processes each of which makes use of the activity of those upstream. The same applies, though more controversially, in the temporal dimension. One way for memory to operate, and a way that makes it a clear case of the sender-receiver set-up, is for it to have a *write-read* structure. This is seen in a Turing machine, for example, where marks are first made on, and later read from, a tape. This is also familiar from the use of external mnemonic tools such as notes and diaries. Many neuroscientists make a point of claiming that this is *not* how memory works in the brain, however. Christof Koch (1999) says: "memory is everywhere, intermixed with computational elements." Some disagree with this; Randy Gallistel and Adam King (2010) have argued that given what brains can evidently do,

---

[5] See O'Connor's paper “Evolving to Categorize.”

[6] Exceptions are discussed in Godfrey-Smith (forthcoming).

they must contain an "addressable read/write memory," one that is more like computer memory. John Donohoe, a defender of the orthodox view, wrote a review of Gallistel and King in which he said that what has been learned is that the brain in fact contains a "write-only" memory system (2010). Gallistel replied that if the brain wrote something but could not read it, that would surely be pointless. Donohoe's view might be better expressed by saying that what we have learned – according to mainstream views – is that the brain contains a *write-activate* memory. Evolved neural machinery has the function of introducing marks or traces into the brain as a result of experience, where these marks usefully affect later behavior, but the marks have these effects without being *read*. From the point of view of the sender-receiver model, this is a debate about whether neural memory is a clear realization of a sender-receiver structure or a more partial, washed-out one, because the entanglement of the roles of sign and sign user.

4. Genetic Memory

In the rest of this paper I will apply these ideas to a special case. Informational and communication-related concepts have become important parts of modern biology, especially the description of gene action. The status of these descriptions has been much debated. Some recent work has cast the problem within a sender-receiver model, or a relative of it. I have in mind papers by Nicholas Shea (2007, 2012) and by Carl Bergstrom and Martin Rosvall (2009). I have discussed their proposals elsewhere.[7] Here I will sketch a different view.

The best justification for a view of gene action that takes concepts of coding and representation (etc.) seriously is the reality of the reader step. Genes have a rather clear reader mechanism, the transcriptional and translational machinery. The rest of the story is not so clear. One way the informational view of genes has been discussed from time to time, mixed in with other ways, is with the idea of memory. David Nanney, in 1958, used an analogy with a library. He said there are two tasks that cells manage to perform:

---

[7] For discussion of Shea see Godfrey-Smith and Sterelny (2007), for discussion of Bergstrom and Rosvall see Godfrey-Smith (2011). For further discussion of this problem see Griffiths (2000) and Godfrey-Smith (2000),

> On the one hand, the maintenance of a "library of specificities," both expressed and unexpressed, is accomplished by a template replicating mechanism. On the other hand, auxiliary mechanisms with different principles of operation are involved in determining which specificities are to be expressed in any particular cell. … [These] will be referred to as "genetic systems" and "epigenetic systems." (1958, p. 712)

I think this is a helpful way of looking at the situation, now as well as then. A genome can be seen as containing a memory of the structure of useful protein molecules. This is not the only thing a genome does, as it also participates in the regulation of its own expression, but the memory role is one role. Specifically, though, the genome is a *cell*-level memory. The reading process that the genome is involved with is an intra-cellular process, and the result of the reading is production of a protein molecule. Cells read their genes, and cells persist and divide, carrying their internal library with them. Organisms are the results, the upshots, of what the cells are doing.

If that is the reader, who is the writer? Which sender's activities make it the case that the cells contain DNA sequences worth reading? The short answer is that there is no sender.

There is a temptation to look automatically for a sender as complement to any receiver. As noted earlier in this paper, that is not always how things are. Unsent signs or cues can be useable. But in general, I said, without a choosing, evolving, or learning sender there is no way for a complicated and elaborate message worth reading to arise. More exactly, the relationship that is relevant here is one between sign complexity and reader complexity. All events leave traces that a smart reader can make use of by inference, and this can include traces that are very complicated. For simpler, mechanical readers, however, the only unsent "messages" (scare-quotes because these are really just cues) worth reading are simple ones. Turning again to genes, we find that the system of gene expression involves a simple, mechanical reader, and what it reads is complicated. So, Shea and Bergstrom and Rosvall might then argue, the good design and the the amount of useful information in DNA shows that it was *made for* an information-carrying role and its messages do have a sender. Perhaps the sender is the previous generation.

But DNA is a special case, which falls outside the application of this argument, because of its place within the evolutionary process. DNA's evolutionary embedding makes it an unsent sign, something worth reading without having being written, even though it is very complex.

The evolutionary process shapes the genetic sequences that are around at any time. Evolution itself is not a sender, it is just something that happens. Mutations happen; differential survival and reproduction happen. As a result of how they happen, DNA sequences are made into libraries of useful specificities, and also into components in control systems that regulate how the library is accessed. Rather than a *write-read* memory system, genetic memory systems are an *evolve-read* system. The *retention* of information, especially across cell division, is an adapted process, but its *inscription* is not.[8]

If so, the contrasts with a paradigm sender-receiver system that are seen in psychological memory – according to mainstream neuroscience – and in genes, can be represented as in Figure 3:

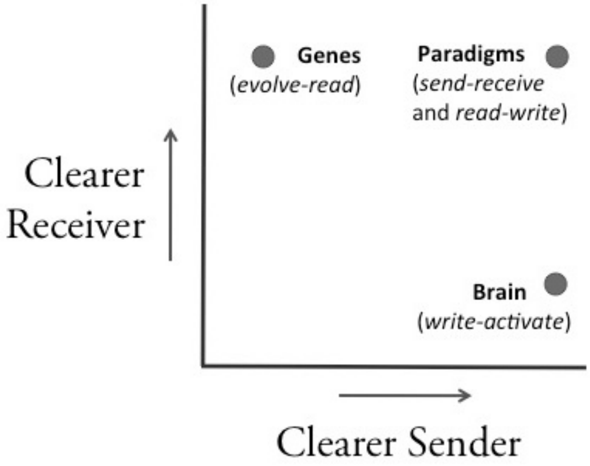


Figure 3: Contrasts between the paradigm sender-receiver set-up and two empirical cases

If this view is right, DNA does not contain a message of the kind seen in signs that have co-adapted senders and receivers. It is a special case of a "cue," in the

---

[8] In contrast, *epi*genetic systems, which function as shorter-term memory, do have producers, inscribers, senders, as well as evolved mechanisms for retention. I hope to discuss this comparison in detail in another paper.

biologist's sense, an unsent sign. But it is still possible to ask what is being read *from* it; what sort of involvement of the DNA sequence with something external to it makes the "message" worth reading? Something like the same distinction between fact-tracking and purely coordinative roles for signs, introduced in Section 2, can be applied here again, and some interesting questions arise which I will only briefly sketch. Dawkins (1998) suggests that our genomes contain a description or model of ancestral worlds. He compares the long-term population-wide representation of the world in our genes with the short-term one that each individual generates in their brain as they live their life. If I am right, then although idea of memory is indeed applicable in both contexts, the details differ from Dawkins' account on two counts. First, there is no genetic representation that is made, sent, or inscribed. The evolutionary process gives rise to genomes worth reading, but no one writes them down. Second, the kind of message-like features that the genome has are not analogous to a representation of how things are, but a representation of what to make, what to do. This is akin to the purely coordinative role for signaling, more like the remembering of a routine or a song, remembering what to make and how to be. There are connections here to analyses of the content of DNA messages that see this content as *imperative* rather than indicative (Stegmann 2005), and also, given the unwritten nature of the genetic message, to "fictionalist" analyses of genetic information (Levy 2011).

* * *

References

Bergstrom, C. T., and Rosvall, M. (2010). The Transmission Sense of Information. *Biology & Philosophy* 26: 159–176.

Dawkins, R. (1998). *Unweaving the Rainbow: Science, Delusion and the Appetite for Wonder*. Boston: Houghton Mifflin.

Donohoe, J. (2010). Man as Machine: A Review of *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*, by C.R. Gallistel and A.P. King. *Behavior and Philosophy* 38: 83-101.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

Gallistel, C. R., & King, A. P. (2009). *Memory and the Computational Brain: Why Cognitive Science will Transform Neuroscience*. New York: Wiley-Blackwell.

Gallistel, C. R. (2010). Response to Donahoe Review. *Behavior and Philosophy* 38: 103-111.

Godfrey-Smith, P. (2000). On the Theoretical Role of 'Genetic Coding.' *Philosophy of Science* 67: 26-44

Godfrey-Smith, P. (2009). *Darwinian Populations and Natural Selection*. Oxford: Oxford University Press.

Godfrey-Smith, P. (2011). "Senders, Receivers, and Genetic Information: Comments on Bergstrom and Rosvall," *Biology and Philosophy* 26: 171-181.

Godfrey-Smith (forthcoming). Evolution and Memory.

Godfrey-Smith, P. and K. Sterelny (2007). "Biological Information," *The Stanford Encyclopedia of Philosophy (Winter 2007 Edition)*, E. N. Zalta (ed.), <http://plato.stanford.edu/entries/information-biological/>.

Griffiths, P. E. (2001). Genetic Information: A Metaphor in Search of a Theory. *Philosophy of Science* 68: 394 – 412.

Harms, W. F. (2004). Primitive Content, Translation, and the Emergence of Meaning in Animal Communication. In D. K. Oller & U. Griebel (Eds.), *Evolution of Communication Systems: A Comparative Approach* (pp. 31–48). Cambridge: MIT Press.

Huttegger, S., Skyrms, B. Smead. R. and Zollman, K. (2010). Evolutionary Dynamics of Lewis Signaling Games. *Synthese* 172: 177-191.

Levy, A. (2011). Information in Biology: A Fictionalist Account. *Nous* 45: 640–657.

Lewis, D. K. (1969). *Convention: A Philosophical Study*. Cambridge, MA: Harvard University Press.

Millikan, R. G. (1984). *Language, Thought and Other Biological Categories*. Cambridge MA: MIT Press.

Nanney, D. L. (1958). Epigenetic Control Systems. *Proceedings of the National Academy of Sciences*, *44*, 712.

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *27*, 379 – 423, 623 – 656.

Shea, N. (2007). Representation in the Genome, and in Other Inheritance Systems. *Biology and Philosophy* 22: 313–331.

Shea, N. (2012). Two Modes of Transgenerational Information Transmission. To appear in *Signaling, Commitment, and Emotion* (ed. R. Joyce, K. Sterelny, and B. Calcott). Cambridge MA: MIT Press.

Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge, MA: Cambridge University Press.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure.* Cambridge, MA: Cambridge University Press.

Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. New York: Oxford University Press.

Stegmann, U. (2005). Genetic Information as Instructional Content. *Philosophy of Science* 72:425-443

Wagner, E. (2011). Deterministic Chaos and the Evolution of Meaning. *British Journal for the Philosophy of Science*. Online first at: doi: 10.1093/bjps/axr039