Draft Dec 1, 2007 Brian Skyrms

Pragmatics, Logic and Information Processing

LOGIC

David Lewis wrote *Convention* in order to use game theory to answer Williard van Orman Quine's skeptical doubts about the conventionality of meaning. Quine's skepticism was directed at the logical positivists' conventional theory of meaning in the service of a conventional theory of logic. According to the logical positivists, logical truths were true and logical arguments valid by virtue of the meanings of the terms involved.

In a famous essay, "Truth by Convention," Quine argued that Positivist accounts of convention (by explicit or implicit definition) required the pre-existence of logic. David Lewis replied that the existence of a convention can be thought of in a different way, as a strong kind of equilibrium or near-equilibrium in a signaling game played by a population. Lewis did not himself supply an account of how a population might get to signaling-system equilibrium – he did not think he had to in order to answer Quine – but the dynamics of evolution or of learning may supply such an account.

Consider the simplest sender-receiver signaling game, of the kind introduced by Lewis in his book *Convention.* One player, the *sender*, observes which of two possible states of the world is actual. (Nature chooses the state with the flip of a fair coin.) The sender chooses one of two signals to transmit to the other player, the *receiver*. The receiver observes the signal, and chooses one of two possible acts. The sender and receiver have common interest in information transmission. Each gets a payoff of one if the right act is chosen for the state, and a payoff of zero otherwise. The signals have no prearranged meanings Any meaning that they acquire is a function of the information transmitted in an equilibrium of the game. An equilibrium in which there is perfect information transmission, where the players always get it right is a *signaling system.* There are two possible signaling systems in this game. In one signaling system the signals have just the opposite meaning that they do in the other, so we have a model in which meaning is perfectly conventional. (There are other equilibria where no information is transmitted.)

A positive answer to the dynamical question raised by Lewis model of convention would show how the dynamics could, at least sometimes, spontaneously generate meaning. For this simplest example, there are strong affirmative answers. *Evolution*, as embodied by the replicator dynamics, always generates a signaling system. (Cite Huttegger) Naive *reinforcement learning*, in the form of Herrnstein's matching law, always generates a signaling system. (Cite Argiento et al here). In more complicated sender-receiver games, perfect signaling can still be generated spontaneously although it is no longer guaranteed to happen with probability one.

These accounts leaves open the question of whether the account thus established in any way supports the doctrine of the conventionality of logic, whose refutation was

Quine's ultimate goal. In so far as one's view of logic is expansive – some positivists viewed all of mathematics as logic – the project may seem Quixotic. But the more modest goal of seeing whether we can get some logic out of information transfer in sender-receiver games, and if so how much, is something worth pursuing.

In previous work[1], I advanced some tentative suggestions that involve modifications to the basic Lewis signaling setup. First, the sender may not observe the state of the world exactly. Rather her observation may rule out some possibilities while leaving a class of others viable. For example, a Vervet monkey may detect the presence of a ground predator - leopard or snake – without being able to see which it is. If this happens often enough and if, as is quite possible, the receiver's optimal action given this information is different from both the optimal action for a leopard or for a snake, it is plausible that a special signal could evolve for this sender's information state in exactly the same way as in the original signaling games. I call such a signal *proto truth-functional* because one way of giving its meaning is by the truth function "leopard or snake" – even though the signal itself is just a one-word sentence. Let us postulate a rich signaling environment in which lots of proto-truth functional signals have evolved.

The second modification is to consider multiple senders, each in possession of a different piece of relevant information. For example, suppose one sender on the ground – seeing a movement of grass - sends the proto-truth function "leopard or snake," and another sender, from the vantage point of a tree, sends the proto-truth function "no leopard." (Negative signals that cancel part of an alarm call are not unknown in animal signaling, as we saw in chapter 2.) Selection favors the receiver who takes the evasive action appropriate for a snake. Such a receiver has performed – or acts as if she has performed - logical inference.

This story was put forward in a tentative and preliminary spirit and it leaves questions hanging. The proto truth-functions were assumed to have already evolved. Could they co-evolve with logical inference, or are they required to exist already? Where are the precise models? Where is their analysis in terms of evolutionary or learning dynamics? We would like to make some progress towards answering these questions, but first we need to show that considerations other than just logic are on the table.

CONVERSATION

In 1967-68, H. Paul Grice delivered seven William James Lectures at Harvard University on *Logic and Conversation*. They were not published in their entirety until after his death, but the lectures themselves had a big impact in Cambridge. The text of the lectures circulated widely as thick typescripts. These lectures played a large part in reawakening linguists to pragmatics, which had been previously been eclipsed by issues of syntax.

Grice was interested in how information over and above conventional meaning was transferred in conversation. Grice's unifying idea was that conversation is fundamentally a cooperative enterprise, and that the presumption of cooperative intent can be used to extract information. If you tell me you have run out of gas and I say that there is a gas station around the corner, you can presume that the gas station is open or at least that I do not know it to be closed. That is so even though my statement would be literally true in either case. You assume that I am trying to cooperate, and that requires truth but also more than mere truth. If I ask you where Peter is, and you answer that he is either in Mexico or Zimbabwe, I can presume that you are not saying this simply because you know that he is in Mexico, although that would make it true. If you are trying to cooperate and know that Peter is in Mexico, you will say so.

Grice and followers in this tradition derived various norms of conversation from the presumption of cooperation, which itself was ultimately elaborated in terms of common knowledge of cooperative intentions. He called these norms "maxims"; we have seen the most important ones – the maxims of quality and quantity – at work in the previous examples.

Quality:

1. Do not say what you believe to be false.
2. Do not say that for which you lack adequate evidence.

Quantity:

1. Make your contribution as informative as required.
2. Do not make your contribution more informative than is required.

In adversarial proceedings, where cooperation cannot be presumed, the oath taken by witnesses is essentially to follow these maxims.

In the basic Lewis signaling games, the foundation of cooperation is made explicit. It is the strong common interest assumed in the specification of the payoffs. The sender is reinforced if and only if the receiver is. This assumption specifies an environment most favorable to evolution of signaling systems. This is not to say that some information transmission may not be possible in the absence of complete common interest – a point to which we will return later.

The presumption of common knowledge that we find in Gricean theory is for the first time made explicit, in David Lewis' account of conventions in signaling games. Although Lewis' theory has affinities with Grice, his focus was quite different. Lewis was giving an account of conventional meaning, while Grice was interested in analyzing information transmission outside the bounds of conventional meaning. Since Lewis is talking about conventions rather than implicatures, the common knowledge is about social patterns of behavior and about the expectations of members of the relevant social group rather than about another participant in a conversation. Conventions, however, can be thought of as having crystallized out of the interplay of the pragmatics of individual conversations.

In our treatment intentions, expectations, common knowledge, and rational choice are not required. They may very well be part of the story when they are present, but there is still a story when they are not. Adaptive dynamics takes center stage. Strongly stable equilibria take the place of common knowledge. This setting is – I believe – a congenial setting for the radical pragmatic theory initiated by Grice and Lewis.

LOGIC and CONVERSATION

When I was first telling my story about evolution of logic, Anil Gupta took me aside after a talk and said: “It’s not just logic that’s evolving, its logic together with conversational implicature and related pragmatic considerations. They only get separated later on in the evolution of language.” Barry Lam said the same thing quite independently and at about the same time in a paper written for me when he was an undergraduate.

Surely they are right. My senders put are all that they know into the “premises” that they are respectively sending. It is just that they don’t know everything. Their knowledge is limited by their imperfect observation of the state of the world. And my receiver is not performing – or acting as if he were performing – an arbitrary valid inference from the premises. Rather he is drawing the most specific valid inference from the premises, which is what he needs to take the optimal action. If this isn’t all exactly Grice, it is consonant with the spirit of Grice. Different maxims resonate with different parts of story.

On the other hand, it is immaterial to the account that the senders communicate everything that they observe. Even if each sender knew that the predator was a snake, and each, for whatever reason, chose to withhold some of the information, the account would work in the same way. It is important here that the senders “tell the truth” – that a sender not send a “snake or leopard” signal while watching an eagle descending from the sky. Senders here need to obey the maxims of quality but not of quantity. And what is important for the receiver is that he extracts the information from the signals that is relevant to his choice of action.

INFORMATION PROCESSING

It is best to think of our two-sender, one-receiver model as an especially simple case of a problem of *information processing*. Multiple senders send signals that convey different pieces of information and the receiver can benefit from integrating this information. Let us consider some simple examples.

1. Inventing the Code:

Suppose that there are four equiprobable states of nature, and that two individuals are situated to make incomplete observations of the state. The first sees whether it is in {S1, S2} or in {S3, S4} and the second sees whether it is in {S1, S3} or in {S2, S4}. Together they have enough information to pin down the state of nature, but separately they do not. Each sends one of two signals to a receiver who must choose one of four acts. The payoffs favor cooperation. Exactly one act is “right” for each of the states in that each of the individuals is reinforced just in case the “right” act for the state is chosen.

I will not assume here, as I did in the story at the beginning of this chapter, that a convention has already been established for the signals used by the senders. We will make things a little harder and *require that the content of the signals evolve together with the inference*. You could think of sender one as waving either a red or a green flag and sender two as waving either a yellow or a blue one.[2]

A *signaling system* in this extended Lewis signaling game is a combination of strategies of the three players, two senders and one receiver, such that the receiver always does the right thing for the state. If we run simulations of reinforcement learning, starting with everyone out acting at random, the three individuals typically fall rapidly into one of the possible signaling systems.

Consider the flow of information in these signaling-systems. Each sender's signal conveys perfect information about her observation – about the partition of states of the world that she can see. The combination of signals has perfect information about the states of the world. Exactly one state corresponds to each combination of signals. And the receiver puts the signals together. The receiver's acts contain perfect information about the state of the world.

2. Inventing the Categories and the Code:

In the foregoing example, we postulated the categories that the senders can observe and thus those that could be embodied in their signals. For example, sender one can at best convey the information is in one of the first two states or not. That is all that she can see. In a remarkable analysis, Jeffrey Barrett considers a model[3] where the two senders and one receiver need to interact to spontaneously invent both the categories and the code in order to achieve a signaling system.

In Barrett's game there are four states and four acts, just as before, but each sender can observe exactly the true state of the world. Although each sender now has perfect information, each only has two signals available. There are two information bottlenecks. To achieve a signaling system our three individuals face a daunting task. Senders need to attach their signals to categories in such a way that these categories complement each other and jointly determine the state of the world. The receiver needs to extract the information from these signals. Receivers need to learn while senders are learning how to categorize and senders need to learn their complementary categorizations while receivers are learning to extract information from the combination of signals received.

In a signaling system, sender 1 might send her first signal in states 1 and 2 and her second signal otherwise, and sender 2 might send her first signal in states 1 and 3 and her second otherwise. (These are just the categories imposed by observational restrictions in example 1.) But alternatively sender 1 might lump states 1 and 4 together for one signal and states 2 and 3 for another which, together with the same receiver's strategy, would let the combination of signals peg the state of the world.

To my considerable surprise, Barrett found that reinforcement learners reliably learned to optimally categorize and signal. The categories formed depended on the vicissitudes of chance – sometimes one set, sometimes another – but they always complemented one another in a way that allowed the receiver to do the right thing.

Consider the flow of information in the signaling-system equilibria in Barrett's game. Sender's signals do not convey perfect information about their observations, but only partial information. Nevertheless, the combination of signals has perfect information about the states of the world. Exactly one state corresponds to each combination of signals. And the receiver puts the signals together. The receiver's acts contain perfect information about the state of the world.

Senders and receivers have, in a way, learned to obey the maxims of quality and quantity while learning to communicate. The receiver has also, in a way, learned to perform a most basic logical inference: from premises *p, q* to infer the conjunction *p & q*.

3. Extracting Relevant Information

Appropriate information processing depends on the character of the payoffs. Let us revisit example 1. The two senders again have their categories fixed by observation. Sender 1 can see whether the world is in one of the first two states or not; sender 2 can see whether the state is odd numbered or even numbered. We modify the example so that there are only two relevant acts with the following payoffs:

	Act 1	Act 2
State 1	0	1
State 2	1	0
State 3	1	0
State 4	0	1

Optimal signaling requires the receiver to do act 1 in states 2 and 3 and act 2 otherwise. Although there are only two acts now, the receiver cannot rely on only one sender, since neither has the sufficient information. The senders have information about their own categories – their own partitions of the states of the world - but the receiver needs information about a different partition. Reinforcement learners, starting with random exploration, learn optimal signaling here just as well and just as quickly as in the previous examples.

Given optimal signaling, where players are always reinforced, each sender's signal here carries perfect information about her observation and the combination of signals singles out the state of the world. But the receiver's *act* only contains partial information about the state. It is "only as informative as is required" by the pragmatic considerations embodied in the reinforcement structure. The receiver has learned to extract the information that is relevant and to ignore that which is irrelevant.

From the viewpoint of truth-functional logic, the receiver has had to learn how to compute the truth-value of the exclusive disjunction, "*nor*", from the truth values of its constituents. Sender 1 observes whether *p* is true; sender 2 observes whether *q* is true. The act that pays off is act 1 if *p nor q*, act 2 if not.

If we look at this in terms of logical inference we can say that the receiver has – in a way – learned to infer *p nor q* from the premises *p, not-q*, but its denial from the premises *p*, *q*, and so forth. The inferences are not just valid inferences, but also the *relevant* valid inferences for the task at hand. Receivers can learn to compute other truth functions and to perform other inferences in just the same way.

4. Taking a Vote

So far, our senders have been infallible observers of the state of the world. They may not have seen everything, but what they think they have seen they have indeed seen. Senders' strategies so far have been based on the truth, of not always the whole truth. In the real world there is observational error[4].

If there is imperfect observation, it may make sense to ask for a second or third opinion. Consider the most basic Lewis signaling game, with two equiprobable states, two signals and two acts, but with three senders. Each sender observes the state, but with some error – errors independent – and sends a signal to the receiver. Then the receiver chooses an act.

It is not possible for signals to carry perfect information about the state. Error is endemic to the model. It is not possible for a signaling system to assure that the receiver always gets it right. But it is possible for an equilibrium to minimize the effects of error. The senders can convey perfect information about their fallible observations, and the receiver can pool this information to make the best choice. The optimal receiver's strategy is then to take a vote. If the majority of senders "say" it is state one, then the receiver should do act one; if a majority of senders "say" it is state 2 then the receiver should do act 2. We could call this sort of equilibrium a "Condorcet signaling system." Taking a vote allows a significant improvement over the payoffs attainable with only one sender. For example, with an error rate for observations of 10%, out receiver will have an error rate of less than 3%. With a few more senders the error the error rate can be driven very low, as the Marquis de Condorcet pointed out.

BEYOND COMMON INTEREST

The strong common interest in Lewis signaling games is by design; they were conceived as situations conducive to the existence of effective and stable norms of communication. Often signals are exchanged in completely cooperative situations, but not always, and the study of signaling should not be restricted to situations of common interest.

Consider a one-sender, two-state, three-act signaling game with conflicting interests (sender's payoffs first, receiver's second):

	Act 1	Act 2	Act 3
State 1	3, 10	0, 0	11, 8
State 2	0, 0	1, 10	10, 8

If the receiver knows the state of nature, he will be best off taking act 1 (in state 1) or act 2 (in state 2), for a payoff of 10. If he is reasonably uncertain, he is better off taking act 3, which has a good payoff (8) irrespective of the state. But the sender wants the receiver to take act 3, which gives her the best payoff in each state. Accordingly, it is in the sender's interest to refrain from communicating information about the state. No signaling system is an equilibrium. But there are pooling equilibria where the sender always sends the same signal and the receiver always does act 3.

Let us add another state where there *is* common interest. In the following example it is very important for both sender and receiver that act 2 be done in state 3.

	Act 1	Act 2	Act 3
State 1	3, 10	0, 0	11,8
State 2	0, 0	1, 10	10,8
State 3	0, 0	10, 10	0,0

Now we have a partial pooling equilibrium, where the sender sends one signal in states 1 and 2, and another signal in state 3 and where the receiver does act 3 on getting the first signal and act 2 on getting the second. In this situation there is a mixture of common and opposed interests. There is an equilibrium in which some information is transmitted from sender to receiver, but not all.

We have a new example where one of the signals comes to have the content “state 1 or state 2.” Here this is not because the sender is unable to discern the state, but rather because she is unwilling to transmit specific information. We have information transmission even though the cooperative principle and the maxim of quantity are violated.

A receiver who gets the signal with the content “state 1 or state 2” might be able to combine this information with that in a signal from a different sender with different payoffs, as in our original example of logical inference.

DECEPTION

Fireflies use their light for sexual signaling. In the western hemisphere, males fly over meadows flashing a signal. If a female on the ground gives the proper sort of answering flashes, the male descends and they mate. The flashing “code” is species-specific. Females and males in general only use and respond to the pattern of flashes of their own species.

There is, however, an exception. A female firefly of the genus Photuris, when she observe a male of the genus Photinus, may mimic the female signals of the male’s species, lure him in, and eat him. She gets not only a nice meal, but also some useful protective chemicals that she cannot get in any other way. One species, Photuris versicolor, is a remarkably accomplished mimic – capable of sending the appropriate flash patterns of eleven Photinus species.

I would say that this qualifies as deception, wouldn't you? Intentionality doesn't really come into it. It is not intellectually profitable to imagine oneself inside the mind of a firefly. Deception is a matter of information. The Photinus species have their signaling systems in place and encounters with Photuris are not sufficiently frequent to destroy it. More precisely, the signals carry information – not perfect information, but information nonetheless – to the effect that state of the world is "receptive female here." This is all a question of frequencies, not intentions. Without the signal it is very improbable that there is a receptive female; with the signal it is probable overall that there is one. Photuris species systematically send signals that carry misleading information about the state of the world. As a consequence the receiving males are led to actions that benefit the senders, but lead to their own demise.

How is deception possible? It is possible because signals need not be perfectly reliable in order to carry information - and because some level of deception is compatible with the existence of informative signaling,[5] just as some level of error is compatible with informative signaling.

INFORMATION PROCESSING and the COOPERATIVE PRINCIPLE

The simplest Lewis signaling games specify an optimal environment for the evolution of information transfer by signals with conventional meaning. The cooperative principle is built into the model with the assumption of strong common interest – senders and receivers both paid off if and only if the receiver does the appropriate act for the state. A signaling system is the realization of optimal cooperation. The sender always knows the state of the world, and has no reason to withhold information or to engage in deception. In a signaling system the receiver need not perform any inference, but only has to pay attention to the state of the world indicated in the signal and do the act most beneficial to her.

The optimal environment for signaling is not the only environment in which signaling can spontaneously arise. It is perhaps the first case to investigate, but it is not really the most interesting case. We have looked at a few simple generalizations that relax the idealizations of the optimal environment. Information transfer can occur when there are partially conflicting interests, where there are errors, and where there is deception. It can be arise where the receiver has to combine the partial information in distinct signals from different sources. This might involve making a logical inference or taking a vote. Information processing may indeed be a way of making up for failures of the cooperative principle, of observation, and in general of the idealized optimal model.

What is left of the cooperative principle and of Grice's maxims? They have a role to play in certain situations, but none of them are required for information transmission and processing. What *is* essential is contained in the pragmatics of sender-receiver interaction.

REFERENCES

Barrett, Jeffrey A. (2006) “Numerical Simulations of the Lewis Signaling Game: Learning Strategies, Pooling Equilibria, and the Evolution of Grammar” Working Paper MBS06-09. University of California, Irvine.

Barrett, Jeffrey A. (2007a) “The Evolution of Coding in Signaling Games” *Theory and Decision*.

Barrett, Jeffrey A. (2007b) “Dynamic Partitioning and the Conventionality of Kinds” *Philosophy of Science* 74:527-546.

Bergstrom, Carl T. and Michael Lachmann (1998) “Signaling Among Relatives III. Talk is Cheap” *Proceeding of the National Academy of Sciences USA* 95:5200-5105.

Grice, H. Paul (1957) “Meaning” *Philosophical Review* 66: 377-388.

Grice, H. Paul (1975) “Logic and Conversation” In *Syntax and Semantics, v.3* Ed. P. Cole and J. L. Morgan. New York: Academic Press, 41-58.

Grice, H. Paul (1989) *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.

Nowak, M. A. and D. Krakauer (1999) “The Evolution of Language” *Proceedings of the National Academy of Sciences of the USA* 96: 8028-8033.

Searcy, William A. and Nowicki, Stephen (2005) The Evolution of Animal Communication: Reliability and Deception in Signaling Systems Princeton: Princeton University Press.

Skyrms, Brian (2000) “Evolution of Inference” In *Dynamics of Human and Primate Societies* ed. Tim Kohler and George Gumerman. New York: Oxford University Press 77-88.

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

NOTES

[1] Skyrms (2000), (2004).

[2] The receiver will have to be able differentiate information from the two senders, since they are placed to make different observations.

[3] Barrett (2006)(2007a,b).

[4] Nowak and Krakauer (1999) consider a different kind of error, receiver's error in perceiving the signal. They suggest that minimizing this kind of error played an important role in the evolution of syntax.

[5] For other examples where deception is compatible with informative signaling, see Bergstrom and Lachmann(1998); Searcy and Nowicki (2005).