Theory and Decision (2009) 67:223–237
DOI 10.1007/s11238-007-9064-0

JEFFREY A. BARRETT

THE EVOLUTION OF CODING IN SIGNALING GAMES

ABSTRACT. Signaling games with reinforcement learning have been used to model the evolution of term languages (Lewis 1969, *Convention*. Cambridge, MA: Harvard University Press; Skyrms 2006, "Signals" Presidential Address. Philosophy of Science Association for PSA). In this article, syntactic games, extensions of David Lewis's original sender–receiver game, are used to illustrate how a language that exploits available syntactic structure might evolve to code for states of the world. The evolution of a language occurs in the context of available vocabulary and syntax—the role played by each component is compared in the context of simple reinforcement learning.

KEY WORDS: evolution of language, evolutionary game theory, signaling games

1. LEWIS SIGNALING GAMES

Lewis (1969) introduced sender–receiver games as a way of investigating how meaningful language might evolve from initially random signals. A Lewis signaling game has two players: the sender and the receiver. In an n-state/n-term signaling game there are n possible states of the world, n possible terms the sender might use as signals, and n possible receiver actions, each of which corresponds to a state of the world. Nature chooses a state at random on each play of the game. The sender then observes the state and sends a term to the receiver, who cannot directly observe the state of the world. The receiver chooses an act based on the term he receives. If the receiver's action matches the state of the world, then each player is rewarded.

The sender and receiver may learn from their record of success and failure on repeated plays of the game. Whether and how quickly they learn will depend on their learning strategy. If the sender and the receiver evolve to a state where they are more successful then chance, then they have evolved a more or less efficient language. The efficiency of the evolved language can be measured by the expected signal success rate (the expected ratio of successful actions to the number of plays) or by the mean information content of a signal (where $\log_2(n)$ bits is sufficient to specify a particular state from among n possible states). Lewis called a system that evolves to a maximally efficient language a *signaling system*. For a perfect signaling system in a Lewis signaling game, each state of the world corresponds to a term in the language, and each term corresponds to an act that matches the state of the world; consequently, each signal leads to a successful action. For the 2-state/2-term game, a perfect signaling system would have a success rate of 1.0 and each signal would communicate one bit of information.

We will begin by supposing the simulated agents use a simple urn learning strategy. Urn learning is a type of positive reinforcement learning with a long psychological pedigree: it models Richard Herrnstein's (1970) *matching law,* where the probability of choosing an action is proportional to the accumulated rewards, which is itself a quantification of Thorndike's *law of effect*. Herrnstein reinforcement learning has been used recently in game-theoretic contexts by Roth and Erev (1995) to model experimental human data on learning in games, by Skyrms and Pemantle (2000) to model social network formation, and by Skyrms (2006) to model learning in the context of Lewis signaling games.

In a basic 2-state/2-term Lewis signaling game with reinforcement learning, there are two possible states of the world (A and B), two possible terms (0 and 1), and two possible acts (A and B), each of which is successful if and only if the corresponding state of the world obtains (see Figure 1). The sender has an urn labeled *state A* and an urn labeled *state B*, and the receiver has an urn labeled *signal 1* and an urn labeled *signal 2*. The sender's urns each begin with one ball labeled *signal*

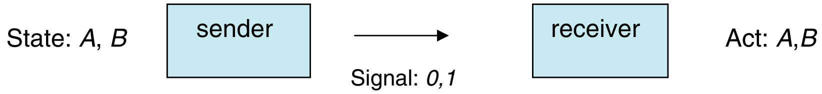

Figure 1.

1 and one ball labeled *signal 2*, and the receiver's urns each begin with one ball labeled *act A* and one ball labeled *act B*.

On each play of the game, the state of the world is randomly determined with uniform probabilities, then the sender consults the sender urn corresponding to the current state and draws a ball at random, where each ball in the urn has the same probability of being drawn. The signal on the drawn ball is sent to the receiver. The receiver then consults the receiver urn corresponding to the signal and draws a ball at random. If the action on the drawn ball matches the current state of the world, then the sender and the receiver each return their drawn ball to the respective urn and add another ball to the urn with the same label as the drawn ball; otherwise, the sender and receiver just return their drawn ball to the respective urn. On this basic urn learning strategy, there is no penalty to the agents for the act failing to match the state. The game is repeated with a new state of the world.

Skyrms (2004, 2006) and Huttegger (2007a,b) have studied 2-state/2-term Lewis signaling games as models for the evolution of term languages. The basic 2-state/2-term signaling game with urn learning is relatively simple and presents a difficult context for the evolution of a successful language. The space of possible states is symmetric with no special saliencies, and a further difficulty is that the learning dynamics is simple reinforcement learning with no penalty for failure. The argument is that if a successful term language can evolve in this context, then it is all the more plausible that a successful language might similarly evolve in contexts with special saliencies or more sophisticated learning strategies.[1]

Skyrms (2006) has shown for urn learning (by simulation) and Huttegger (2007a) has shown for the closely related replicator dynamics (by proof), that perfect signaling always

TABLE I
Run failure rates for Lewis signaling games with urn learning

Model	Run failure rate
3-state/3-term	0.096
4-state/4-term	0.219
8-state/8-term	0.594

evolves in the 2-state/2-term signaling game if states are evenly distributed. Skyrms (2006) has also shown that perfect signaling evolves in a system with two senders and one receiver when the senders observe different, prearranged two-cell partitions of a four-state space.

It is easy to get a sense of how successful evolution of signaling systems works in the 2-state/2-term Lewis signaling game with simple reinforcement learning. Adding balls to the signal and act urns when an act is successful changes the relative proportion of balls in each urn, which changes the conditional probabilities of the sender's signals (conditional on the state) and the receiver's acts (conditional on the signal). The change in the proportion of balls of each type in each urn increases the likelihood that the sender and receiver will draw a type of ball that will lead to successful coordinated action. Here the sender and receiver are simultaneously evolving and learning a meaningful language. That they have done so is reflected in their track-record of successful action.

The situation is more complicated for signaling games with more (or fewer) states or terms or if the distribution of states is biased (see Barrett, 2006; Huttegger, 2007a). In such modified games, partial pooling equilibria may develop and prevent convergence to perfect signaling. Table I shows the run failure rates for Lewis signaling games with more than two states and terms (see Barrett, 2006 for more details). Here there are 10^3 runs of each model with 10^6 plays/run. A run is taken to fail if the signal success rate is less than 0.8 after 10^6 plays.

TABLE II
Distribution of signal success rates in the 8-state/8-term signaling game

Signal success rate interval	Proportion of runs
[0.0, 0.50)	0.000
[0.50, 0.625)	0.001
[0.625, 0.75)	0.045
[0.75, 0.875)	0.548
[0.825, 1.0]	0.406

While these results illustrate failures in uniform convergence to perfect signaling, each system is always observed to do better than chance and hence to evolve a more or less effective language. In those cases where perfect signaling fails to evolve in the 3-state/3-term game, the system nevertheless approaches a signaling success rate of about 2/3.[2] Similarly, in the 4-state/4-term game, when a system does not approach perfect signaling, it approaches a success rate of about 3/4.

The behavior of the 8-state/8-term system is more complicated since there are several partial pooling equilibria corresponding to different signal success rates. The distribution of signal success rates in the 8-state/8-term game with 10^3 runs and 10^6 plays/run is given in Table II.

The partial pooling equilibria that sometimes block convergence to perfect signaling in such games are an artifact of simple reinforcement learning. If one allows for a slightly more sophisticated learning strategy, then one gets a better rate of convergence to perfect signaling. On the 8-state/8-term (+2, −1) signaling game, success is rewarded by adding to the relevant urns two balls of the type that led to success and failure is punished by removing from the relevant urns one ball of the type that led to failure. As illustrated in Table III, this learning strategy more than doubles the chance of perfect sig-

TABLE III
Distribution of signal success rates in the 8-state/8-term (+2, −1) signaling game

Signal success rate interval	Proportion of runs
[0.0, 0.50)	0.000
[0.50, 0.625)	0.000
[0.625, 0.75)	0.002
[0.75, 0.875)	0.110
[0.825, 1.0]	0.888

Figure 2.

naling evolving in the 8-state/8-term game. More sophisticated learning strategies do better yet.[3]

2. SYNTACTIC GAMES AND DYNAMIC PARTITIONING

While the signaling games considered so far illustrate how a simple term language might evolve from random signaling, it is natural to ask about more subtle linguistic conventions. *Syntactic games* are extensions of Lewis signaling games where there are more states relevant to successful action than available terms but also more than one signal is available on each play of the game. The syntactic degrees of freedom provided by multiple ordered signals may then evolve to be used for the representation of states.

In the 4-state/2-term/2-sender syntactic game there are two senders who observe the state of the world, then each sends a signal of either 0 or 1 (see Figure 2). Each signal is independent in the sense that neither sender knows what the other sent. There is one receiver who knows each signal and which sender sent it but does not know the state of the world. There are four acts, each corresponding to one of the four states. A receiver's act is successful if and only if the corresponding state obtains. We will start again with basic urn learning where each of the senders has her own urns and where both senders and the receiver add a ball of the successful signal or act type to the appropriate urn on success and simply replace the drawn balls on failure. We will also suppose uniformly distributed states.

Since the receiver knows which sender sent each signal, the two signals together may be considered to be a single length-two message. The order of the two terms provides syntax that may evolve to represent states.

This is a very difficult context for the evolution of language. Since there are four states and four acts but only two terms, perfect signaling can only evolve if the senders and receiver learn to use the available syntactic structure to code for the state-act pairs. Further, the state space is symmetric with no special saliencies and the simple learning dynamics allows for only positive reinforcement. And since neither sender knows what signal was sent by the other, they cannot directly learn to correlate their signals to code for the state. Nevertheless, the senders and receiver typically evolve a successful language that codes for each of the four state-act pairs.

As with the 4-state/4-term Lewis signaling game, the 4-state/2-term/2-sender syntactic game with basic urn learning typically approaches perfect signaling (approximately 3/4 of the runs are successful). Table IV shows simulation results for the 4-state/2-term/2-sender game with a comparison to the results of the 4-state/4-term game.

On a successful run of the 4-state/2-term/2-sender game, the senders and receiver simultaneously evolve coordinated partitions of the state space and a code where each sender's partial information together selects a state. The code that evolves on a successful run is a permutation of "**00**" means

TABLE IV

Failure rates of the 4-state/2-term/2-sender syntactic game. The 4-state/4-term Lewis signaling game in included for comparison

Number of plays/run	4-state/2-term/2-sender failure rate (< 0.8 signal rate)	4-state/4-term failure rate (< 0.8 signal rate)
10^6	0.269 [2000 runs]	0.219 [1000 runs]
10^7	0.25 [100 runs]	0.17 [100 runs]
10^8	0.27 [300 runs]	0.19 [100 runs]

state 1, "**01**" means state 2, "**10**" means state 3, and "**11**" means state 4. Partial pooling equilibria are responsible for those runs where perfect signaling does not evolve. Such failed runs are observed to approach a signaling success rate of about 3/4, and thus still do better than chance and, in this sense, represent the evolution of a language.

More complex coding schemes evolve in syntactic games with more states, more senders, or more terms. In the 8-state/2-term/3-sender syntactic game, there are eight states and three independent senders, each restricted to the two terms. With basic urn learning, this model approaches perfect signaling about 1/3 of the time. In this case, each of the eight possible states of the world is represented by a length-three binary string. The distribution of signal success rates for 10^3 runs with 10^6 plays/run is given in Table V and the corresponding Lewis signaling game (with one sender and eight terms) is included for comparison. Again, while the evolution of partial pooling equilibria sometimes prevents the evolution of perfect signaling, a more-or-less effective language always evolves.

A slightly more sophisticated learning strategy can significantly improve the chance of evolving perfect signaling in a syntactic game. Table VI gives results for the 8-state/2-term/3-sender (+3, −1) system on 10^3 runs with 10^6 plays/run (the learning strategy is reinforcement where the senders and receiver add three balls of the successful type on success and

TABLE V
Distribution of signal success rates in the 8-state/2-term/3-sender syntactic game

Signal success rate interval	8-state/2-term/3-sender model proportion of runs	8-state/8-term model proportion of runs
[0.0, 0.50)	0.000	0.000
[0.50, 0.625)	0.001	0.001
[0.625, 0.75)	0.081	0.045
[0.75, 0.875)	0.589	0.548
[0.825, 1.0]	0.329	0.406

TABLE VI
Distribution of signal success rates in the 8-state/2-term/3-sender (+3, −1) syntactic game. The distribution for the 8-state/8-term (+3, −1) is included for comparison

Signal success rate interval	8-state/2-term/3-sender (+3, −1) proportion of runs	8-state/8-term (+3, −1) proportion of runs
[0.0, 0.50)	0.000	0.000
[0.50, 0.625)	0.000	0.000
[0.625, 0.75)	0.004	0.004
[0.75, 0.875)	0.157	0.225
[0.825, 1.0]	0.839	0.771

remove one ball of the failed type on failure). Results for the 8-state/8-term (+3, −1) game (one sender with eight terms and with the same positive and negative reinforcement as the syntactic game) is included for comparison. With its upgraded learning dynamics, the 8-state/2-term/3-sender (+3, −1) system approaches perfect signaling on most runs.

TABLE VII

Distribution of signal success rates in the 9-state/3-term/2-sender (+1, −0) and the 9-state/3-term/2-sender (+1.5, −1.0) syntactic games

Signal success rate interval	9-state/3-term/2-sender (+1, −0) proportion of runs	9-state/3-term/2-sender (+1.5, −1.0) proportion of runs
[0.0, 0.6)	0.000	0.000
[0.6, 0.7)	0.005	0.000
[0.7, 0.8)	0.135	0.000
[0.8, 0.9)	0.610	0.051
[0.9, 1.0]	0.250	0.949

Codes involving more than two terms may also evolve in a syntactic game; indeed, the efficiency of simple reinforcement learning in evolving a trinary code is comparable to the efficiency in evolving a binary code for a similar number of state-act pairs. For the 9-state/3-term/2-sender game, there are partial pooling equilibria at signal success rates of 0.88, 0.77, and 0.66. Table VII gives the results of 9-state/3-term/2-sender games on 10^3 runs with 10^6 plays/run with only positive reinforcement and with both positive and negative reinforcement learning. Perfect signaling is approached about 1/4 of the time for positive reinforcement alone and almost always for positive and negative reinforcement together.

In a syntactic game, systematically interrelated partitions of the state space co-evolve with the successful language. There is one partition of the space for each position in the message, a term at a position selects an element in the corresponding partition in such a way that the ordered terms of a message together select a single state at the level of individuation required for successful action. In the 9-state/3-term/two-sender game two coordinated three-cell partitions of the state space evolve so that a length-two message serves to uniquely select a single state.[4]

3. NUMBER OF TERMS, MESSAGE LENGTH, AND THE EFFICACY OF LANGUAGE EVOLUTION

Available vocabulary and syntax may play different roles in the evolution of a language. This can be seen by considering possible evolutionary trade-offs between the number of available terms and available syntax. Is it more difficult, for example, to evolve a successful language with more available terms and a simple syntax or with a more complex available syntax and fewer terms? While one should expect the answer to such a question to be contingent on the particular learning strategy employed, it is possible to get some insight here by assuming simple reinforcement learning, holding constant the number of types of state that must be distinguished for successful action, then varying the available number of terms and the available degrees of syntactic freedom. It is supposed that there is no special cost for using a language that has more terms or a more complex syntax—that is, the agents are assumed to have perfect, cost-free memories and to send and to receive cost-free messages regardless of length.

Suppose that there are 16 types of state that must be distinguished for successful action, and consider three games: (1) a 16-state/16-term/1-sender Lewis signaling game, (2) a 16-state/4-term/2-sender syntactic game, and (3) a 16-state/2-term/4-sender syntactic game. Each of these games has sufficient available terms and/or syntax to evolve perfect signaling if the degrees of freedom evolve to be used optimally.

Table VIII shows the distribution of signal success rates over 10^3 runs for the three games with 10^6 plays/run. Systems with more available terms and less available syntax do slightly better on average than systems with fewer terms and more syntax. The systems with more available terms, however, have a much better chance of successfully approaching perfect signaling in the long run: the 16-state/16-term/1-sender sys-

TABLE VIII

Distribution of signal success rates and mean signal success rates for a 16-state term game and two 16-state syntactic games. 10^3 runs with 10^6 plays/run

Signal success rate interval	16-state/16-term/ 1-sender proportion of runs	16-state/4-term/ 2-sender proportion of runs	16-state/2-term/ 4-sender proportion of runs
[0.0, 0.75)	0.003	0.014	0.017
[0.75, 0.80)	0.067	0.050	0.048
[0.80, 0.85)	0.110	0.131	0.173
[0.85, 0.90)	0.379	0.416	0.439
[0.90, 0.95)	0.369	0.335	0.290
[0.95, 1.0]	0.072	0.054	0.033
Mean signal success rate	**0.8829**	**0.8781**	**0.8704**

tem is more than twice as likely to approach perfect signaling on simple reinforcement learning than the 16-state/2-term/4-sender system.

On the other hand, as illustrated by Table IX, more available syntax and fewer available terms does better than more terms and less syntax in the short run. With 2×10^4 plays/run, the mean signal success rate over 10^3 runs for the 16-state/2-term/4-sender game is higher than for the 16-state/16-term/1-sender game by a factor of about 1.3.

For basic urn learning then more available terms and less available syntax helps in the long-run, and more syntax and fewer terms helps in the short-run. But again, one should, in general, expect such results to be contingent on the particular learning strategy.

4. CONCLUSION

Syntactic games illustrate how language might evolve by exploiting both available vocabulary and available syntax to

TABLE IX

Distribution of signal success rates and mean signal success rates for a 16-state term game and two 16-state syntactic games. 10^3 runs with 2×10^4 plays/run

Signal success rate interval	16-state/16-term/ 1-sender proportion of runs	16-state/4-term/ 2-sender proportion of runs	16-state/2-term/ 4-sender proportion of runs
[0.0, 0.45)	0.530	0.026	0.004
[0.45, 0.50)	0.339	0.164	0.057
[0.50, 0.55)	0.116	0.394	0.223
[0.55, 0.60)	0.015	0.312	0.394
[0.60, 0.65)	0.000	0.100	0.244
[0.65, 1.0]	0.000	0.004	0.078
Mean signal success rate	**0.4485**	**0.5407**	**0.5781**

represent states at a level of description sufficient for successful action. They also show how available vocabulary and syntax might play different roles in the evolution of a successful language. Exactly how they interact with each other should be expected to depend on the particular learning strategy and on any special costs for remembering terms or syntax or for the transmission of longer messages.

ACKNOWLEDGMENTS

I would like to thank Brian Skyrms for many pleasant conversations on the topics at the heart of this article. I would also like to thank Rory Smead, Kevin Zollman, and Simon Huttegger.

NOTES

1. While an even distribution of states may seem to contribute to a difficult environment for language evolution, it is harder for perfect signaling to evolve under simple reinforcement learning when the probability distribution over states of the world is not uniform. The agents might get a good enough success rate by always choosing the more likely state to reinforce the use of more than one term for this state; and since there is no punishment for failure on this learning strategy, there is no evolutionary pressure to undo these reinforced dispositions. See Huttegger (2007a) for more details.
2. Systems that approach a signaling success rate of 2/3 here do not learn to signal reliably with two out of three terms; rather, such systems approach a partial pooling equilibrium where two of the signal terms correspond to the same state-act pair and the other term is used to represent both of the other state-act pairs, and the sender and the receiver follow (different) mixed strategies. See Barrett (2006) for more details.
3. Simulated agents using the Bereby-Meyer and Erev (1998) adjustable reference point with truncation learning model, which was designed to model empirical features of actual human learning, for example, are always observed to approach reliable signaling for some parameter settings of the ARP model (Barrett, 2006).
4. It is a curious feature of all of the signaling games considered in this article that the signal success rate is always observed to be greater than 1/2. While Simon Huttegger has a argument for why the success rate should be better than chance signaling for such games, it is unclear, at least to me, why it should always be better than 1/2. This may be a property related to the sure-fire evolution to perfect signaling in the context of the original two-state Lewis. If so, it may also depend on the even distribution of states.

REFERENCES

Barrett, J.A. (2006), Numerical simulations of the Lewis signaling game: learning strategies, pooling equilibria, and the evolution of grammar, UC Irvine Institute for Mathematical Behavioral Sciences Preprint (22 September 2006). http://www.imbs.uci.edu/tr/abs/2006/mbs06_09.

Bereby-Meyer, Y. and Erev, I. (1998), On learning to become a successful loser: a comparison of alternative abstractions of learning processes in the loss domain, *Journal of Mathematical Psychology* 42, 226–286.

Lewis, D. (1969), *Convention*. Cambridge, MA: Harvard University Press.

Herrnstein, R.J. (1970), On the law of effect, *Journal of the Experimental Analysis of Behavior* 13, 243–266.

Huttegger, S. (2007a), Evolution and the explanation of meaning, *Philosophy of Science*.

Huttegger, S. (2007b), Evolutionary explanations of indicatives and imperatives, *Erkenntnis*.

Roth, A. and Erev, I. (1995), Learning in extensive form games: experimental data and simple dynamical mod3els in the intermediate term, *Games and Economic Behavior* 8, 164–212.

Skyrms, B. (1996), *Evolution of the social contract*. Cambridge: Cambridge University Press.

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

Skyrms, B. (2006), Signals Presidential Address. Philosophy of Science Association for PSA 2006.

Skyrms, B. and Pemantle, R. (2000), A dynamic model of social network formation, *Proceedings of the National Academy of Sciences in the USA* 97, 9340–9346.

Address for correspondence: Jeffrey A. Barrett, Logic and Philosophy of Science, UC Irvine, Irvine, CA 92697-5100, USA. E-mail: jabarret@uci.edu