TRUTH AND PROBABILITY IN EVOLUTIONARY GAMES

JEFFREY A. BARRETT

ABSTRACT. This paper concerns two composite Lewis-Skyrms signaling games. Each consists in a base game that evolves a language descriptive of nature and a metagame that coevolves a language descriptive of the base game and its evolving language. The first composite game shows how a pragmatic notion of truth might coevolve with a simple descriptive language. The second shows how a pragmatic notion of probability might similarly coevolve. Each of these pragmatic notions is characterized by the particular game and role that it comes to play in the game.

1. INTRODUCTION

Simple games may evolve from the ritualization of the decisions of agents, and more complex games may evolve from the composition of simpler games (Barrett and Skyrms 2015). Here we briefly consider two composite Lewis-Skyrms signaling games (Lewis 1969) (Skyrms 2010). Each consists in a base game that evolves a language descriptive of nature and a metagame that *coevolves* a language descriptive of the base game and its evolving language. The first composite game shows how a pragmatic notion of truth might coevolve with a simple descriptive language. The second shows how a pragmatic notion of probability might similarly coevolve.

The primitive notions of *truth* and *probability* are well characterized by the two games and roles that each notion comes to play in the game where it evolves. The thought is that one might use the games where such notions evolve and roles they come to play to individuate between alternative pragmatic understandings of the notions.[1] Here we consider two simple examples of how this might be done. We will begin with a brief discussion of the signaling games and the sort of explanation they provide.

2. SIGNALING GAMES AND EXPLANATION

A simple Lewis-Skyrms signaling game consists of a sender and a receiver. The sender observes a randomly selected state of nature, then sends a signal. The receiver, who does not have direct access to the state of nature, receives the signal then performs an action that either matches the state of nature and is successful

---

[1]See (Barrett 2001 and 2014) for discussions of pragmatic approaches to inquiry where this might fit well.

or does not and is unsuccessful. The agents update their dispositions to signal (conditional on the state of nature) and to act (conditional on the signal) based on the success or failure of the receiver's action.

The sender and receiver may learn to play more successfully on repeated plays of the game by updating their conditional dispositions. The agents' second-order dispositions determine how they update their first-order dispositions to signal and act. Their second-order dispositions also determine what counts as success and failure in action. Indeed, all it means to say that an action is successful in the games we will consider is that it in fact leads to the reinforcement of whatever first-order dispositions produced the action. Similarly, an action is unsuccessful if it does not lead to the reinforcement of the dispositions that produced the action.

Basic reinforcement learning is among the simplest sorts of learning that one might consider.[2] It can be modeled by adding and removing balls from urns, then using draws from the urns to determine one's actions.

Consider a two-state/two-signal/two-act signaling game. Suppose that the sender has one urn for each state of nature labeled 1 and 2, with each urn initially containing a single ball of each of two signal types A and B. And suppose that the receiver has one urn for each signal type labeled A and B, with each urn initially containing a single ball of each of the two act types 1 and 2. On a play of the game, the sender sees the state of nature, draws a ball at random from the corresponding sender urn, then sends the signal indicated on that ball. The receiver sees the signal, draws a ball at random from the corresponding receiver urn, then performs the action indicated on that ball. If the action is successful, each agent returns the ball she drew to the urn from which she drew it, then adds a new ball to that urn of the same type; otherwise, each agent simply returns the ball she drew to the urn from which she drew it.

Initially, such agents will only be successful by blind luck, but as they reinforce on success, they may learn from their experience and perhaps evolve a simple signaling language that allows them to coordinate the receiver's action with the current state of nature. In the two-state/two-signal/two-act case just described, one can prove that the agents will converge to a perfect signaling system with probability one if each state of nature is equally likely.[3] But for the n-state/n-signal/n-act case where $n > 2$ or when nature is biased, the agents may not converge to a perfect signaling system on simple reinforcement learning. When they do not converge to a perfect signaling system, the agents end up in a suboptimal partial pooling equilibria where

---

[2]See Herrnstein (1970) for an early characterization of simple reinforcement learning. More sophisticated forms of reinforcement learning have also been studied. Some of these model human behavior well in some learning contexts Roth and Erev (1995 and 1998) and Barrett and Zollman (2009).

[3]See Argiento, Pemantle, Skyrms, and Volkov (2009).

their evolved first-order dispositions represent mixed conditional strategies on the states and signals.[4]

Successful signaling can evolve in a signaling game faster and more reliably if the agents are equipped with a more sophisticated learning dynamics than simple reinforcement. But if one can show how it is possible for a particular feature of language to evolve by means of a very low-rationality form of learning like simple reinforcement, then the explanation is more generic and hence more compelling. Assuming agents who learn by simple reinforcement does not require one to appeal to special features of the agents that are well-tuned to the particular task at hand. Further, insofar as more sophisticated learning dynamics are typically able to recapitulate the successes of such low-rationality dynamics, showing that a particular feature of language use can evolve on simple reinforcement learning shows that it can evolve on a broad class of stronger learning dynamics. Finally, among low-rationality learning dynamics, simple reinforcement has the virtue that the evolution of the agents' dispositions correspond to the evolution of strategy types in a population on the replicator dynamics, which allows one to translate the evolutionary results from one's learning models to corresponding population models.[5]

Note that the aim is not to explain how the features we find in natural language in fact evolved as that is something we cannot know. Rather, signaling game models explain how it is possible for particular features of language to evolve under generic circumstances and hence without requiring agents with special capacities.

To this end, the fact that the structure of a signaling game is also generic is an important aspect of such explanations. All that is required on such models are systems that are capable of performing actions conditional on states of nature or conditional on the actions of other systems and who may update their first-order dispositions based on what happens. That is, signaling games may evolve from the ritualization of the actions of very simple agents.[6] That basic forms of language may evolve from the interactions of such simple agents illustrates, among other things, that the preconditions for the evolution and use of language are significantly more modest than philosophers have typically thought.[7] And showing how primitive

----

[4]There are other low-rationality types of learning like win-stay/lose-shift where the agents do not get stuck in such suboptimal pooling equilibria. The properties of signaling games on forgetful dynamics like win-stay/lose-shift are a topic for another occasion. See Barrett and Zollman (2009) for an introduction to such learning dynamics.

[5]See Skyrms (2010) for an introduction to learning and population models and the relationship between them and Huttegger, Skyrms, Tarrès, and Wagner (2014) for a more detailed discussion of evolutionary dynamics for signaling games.

[6]See Barrett and Skyrms (2015) for an account of how both simple and composite signaling games may evolve by the ritualization of agent actions.

[7]Concretely, this illustrates that metaphysical notions such as *ideal forms*, *innate ideas*, *pure concepts of the understanding*, or *an a priori language of thought* are not required to explain how

notions of truth and probability might coevolve with descriptive language helps to make the point that even such traditional philosophical language may simply evolve to provide agents with whatever expressive power is in fact required for successful coordinated action given the task at hand.

There are other motivations for modeling the evolution of the notions of truth and probability in particular. To begin, these particular models show how notions for talking about belief and knowledge may themselves evolve, and, hence, provides a step in explaining how simple agents might learn to use notions like truth and probability in the evaluation of evidence and reflection on their own beliefs.[8]

Finally, a concrete evolutionary model explicitly describes a context where notions like *truth* or *probability* might evolve and, consequently, allows one to precisely characterize and individuate alternative pragmatic versions such notions by the roles that they play in particular specified games. When a stable truth or probability predicate evolves in a signaling game, it comes to play a role in coordinating the actions of the agents that that provides them with stable payoffs given whatever their second-order dispositions happen may be. It is in this way that evolutionary models where the predicates evolve allow one to individuate between alternative pragmatic notions of truth and probability. And they are pragmatic notions here insofar as they in fact lead to successful action given the aims of the agents as represented by their second-order dispositions.

3. TRUE AND FALSE

Consider a composite game consisting in a base game and a metagame as represented in Figure 1.[9] The base game is a Lewis-Skyrms signaling game where the sender observes the state of nature. She then sends an initially meaningless signal to a receiver, who does not have access to the state of nature. The receiver then performs an action that is successful if it matches the current state of nature. Both agents are rewarded if the receiver's action is successful. The reward strengthens the dispositions that led to the successful action.

The metagame game is also a Lewis-Skyrms signaling game, but here the sender observes the state of the base game. In particular, she observes whether the sender in the base game sent the signal that she has used most often in the current state of nature. The metagame receiver is rewarded for acting in a way that matches whether the base game sender's signal was in fact her typical signal given the current state of nature.

---

a meaningful language of the sort that facilitates successful coordinated action is possible. Rather, on the present account, even metaphysical-sounding language itself, like language involving *truth* or *probability*, is explained as the result of basic evolutionary processes.

[8]See Barrett (2014) for a general discussion of this point.

[9]This game was first described in Barrett (2015).

If the agents in both games evolve to be uniformly successful, then the base game will have evolved a simple language that is descriptive of nature and the metagame will have evolved a language that is descriptive of the use of base-game language. Indeed, when both games are successful, the base game will evolve basic descriptive expressions and the metagame will coevolve a notion of *truth* that is integrated with the evolved use of these expressions.

We will first consider the base game, then the metagame.

The base game is a $4 \times 4 \times 4$ Lewis-Skyrms signaling game that takes unbiased states of nature as input. On each play, the sender observes one of four states of nature, then sends one of four signals. The receiver, who does not have access to the state of nature, performs one of four actions. The action is successful if and only if it matches the state of nature. If the act is successful, then the disposition that led to each agent's last action is reinforced.

We will assume that the agents' dispositions evolve by simple reinforcement learning. One might imagine the sender with one urn for each state of nature, each of which begins with one ball of each signal type; and the receiver with one urn for each signal, each of which begins with one ball of each act type. On each play, the sender draws a ball at random from the urn corresponding to the current state of nature and sends the corresponding signal; then the receiver draws a ball at random from the urn corresponding to the signal and does the corresponding action. If the action matches the state, then each agent returns the ball to the urn from which it was drawn and adds another ball of the same type. Otherwise, each agent simply returns the ball to the urn from which it was drawn.

On simulation, the agents' dispositions evolve to exhibit a cumulate success rate of better than 0.95 about 0.75 of the time (on 1000 runs with 1×10^6 plays per run). Otherwise, they get stuck in a partial pooling equilibrium and exhibit a cumulative success rate of about 0.75.

The metagame is a $2 \times 2 \times 2$ Lewis-Skyrms signaling game that takes states of nature together with states of the base game as input.[10] In particular, the metagame sender observes whether the base-game sender sends the signal in the current play that she has used most often given the current state of nature. One metagame receiver action is successful if and only if the base-game signal was in fact the most often used in the current state of nature; otherwise, the other is successful. And we will suppose, again, that the agents learn by simple reinforcement.

---

[10]Note that base-game states are not equally probable. Hence, while the metagame is a $2 \times 2 \times 2$ signaling game, and such games do converge on simple reinforcement learning with unbiased nature, the metagame is not guaranteed to converge to a perfect signaling system. See Hofbauer and Huttegger (2008) and Argiento, Pemantle, Skyrms, and Volkov (2009) for detailed discussions of the behavior of $2 \times 2 \times 2$ signaling games.
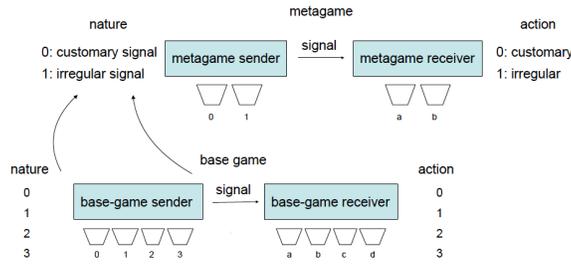
On simulation of the composite game, the metagame sender typically learns to track whether the base-game sender just used whatever has evolved to be the *conventional* signal for a particular state in the base game. In particular, the metagame agents typically (0.78) evolve signals that distinguish between the base-game sender sending the signal that she has sent most of the time in the current state and not (on 1000 runs with 1×10^6 plays per run).

And when they do, they typically also evolve terms that indicate the truth value of the base game signals. Specifically, (1) if the base game evolves a successful signaling system that matches states of nature to successful actions and (2) if the metagame signals coevolve to communicate whether the base-game sender sent the signal that has evolved to be conventional in the base game given the current state of nature, then the metagame signals will reliably indicate whether the base-game sender's signal in fact represents the current state of nature. In this precise sense, then, the metagame terms coevolve to indicate the *truth* of the base-game expressions.

Tracking the most often played signal in the current state requires the metagame sender to exhibit a certain degree of arithmetic competence. Such competence might also evolve by low rationality forms of learning.[11] And if they do, such evolved dispositions might be appropriated for use in the context of a more sophisticated composite game like the one discussed here.[12] The thought is that whatever dispositional capacities the agents have are the product of ritualized actions all the way down.

4. PROBABLE AND IMPROBABLE

A pragmatic notion of *probability* might similarly coevolve with a basic descriptive language if the metagame sender tracks the conditional success rates of actions resulting from the use of the base-game expressions.

---

[11]See Barrett (2013) for an example of how simulated agents may coevolve expressions of cardinalities and the ability to carry out simple arithmetic operations with these expressions using a variety of reinforcement learning.

[12]See Barrett and Skyrms (2015) for a discussion of how this may occur and examples.

Consider the composite game represented in Figure 2. Here the base game is a 2×2×2 Lewis-Skyrms signaling game with unbiased nature and simple reinforcement learning. The sender sees a random state then sends a signal of a or b. Both agents are rewarded if and only if the receiver's action matches the state of nature. In contrast with the base game in the last section, however, we will suppose that the channel between the sender and receiver is noisy. In particular, we will suppose here that, with probability 0.5, the sender's signal is simply set to b on each play of the game.

If a meaningful signaling language evolves in the base game, then the receiver will, consequently, see too many b signals. Hence, the base game agents will at best be successful on 0.75 of their plays.[13]
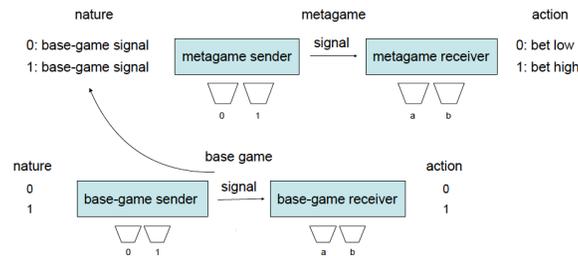
The metagame is a 2×2×2 Lewis-Skyrms signaling game with simple reinforcement learning. But this time the metagame sender tracks whether the current base-game signal, as seen by the base game receiver, has a success rate greater than or equal to a threshold T. To begin, suppose that $T = 0.75$. One metagame receiver action is successful if the relative success rate of the current signal is greater than or equal to T; otherwise, the other action is successful. The metagame receiver might, for example, use the base-game signal to decide whether to accept or reject a wager at odds on the mean reliability of the current base-game signal randomly sampled from the history of past plays.[14]

On simulation of the composite game, the base language typically (0.70) evolves to be as successful as possible given the noise (on 1000 runs with 1×10^6 plays per run).[15] When it does, one base-game signal type evolves to be more reliable than the threshold ($T = 0.75$) and the other less reliable.[16] A metalanguage (that is,

---

[13]This is the case where the base-game agents evolve a reliable signaling system for unbiased nature but about half of the sender's a signals are set to b by the random noise.

[14]We will suppose, for the present, that acting with a positive expected return guarantees success, then weaken this assumption later. And we will also suppose that the metagame sender has previously evolved the arithmetic competence to track success rates.

[15]The best that the agents can do is 0.75, so a cumulative success rate of better than 0.70 was used as the cutoff for success here.

[16]One signal, as seen by the receiver, typically evolves to be reliable about 0.99 of the time and the other about 0.66 of the time.

a language for talking about language itself) that successfully sorts the base-game signals by their reliability above and below T typically (0.73) also evolves.[17]

This assumes that the metagame agents are guaranteed success when the metagame receiver accepts a series of wagers with a positive expected return. The effect of these agents sometimes not being rewarded when the receiver does the right thing, is to lower the metagame success rate as the proportion of such events increase. But, even with just simple reinforcement learning, the metagame typically evolves a language that tracks the relevant probabilistic distinction in the base game for a broad range of ways that one might structure the metagame receiver's wagers.[18]

When a reliable metalanguage does coevolve with the base language, its terms indicate the past conditional reliability of each base-game signal type. So (1) if the past conditional reliability of each base-game signal type evolves to reflect its future conditional reliability and (2) if the metalanguage coevolves to successfully indicate the past reliability of the base game signals, and both conditions are in fact typically met on simulation, then the metagame signals will coevolve to reliably indicate the coarse-grained *probability* of each base-game signal faithfully indicating the current state of nature and hence leading to successful action.

There is a family of such composite probability games indexed by the threshold T. As one varies T, if conditions (1) and (2) are satisfied, then a different coarse-grained distinction evolves. The metagame agents may learn to use these distinctions to express increasingly fine-grained conditional probabilities if they are also able to learn to identify the value of T in the base game they are currently observing.

Finally, note that, in contrast with the truth game above, the metagame here begins by tracking the metagame receiver's coevolving betting behavior rather than the metagame sender's coevolving use of signals. The thought is that what the agents track and how they are rewarded determines the pragmatic content of the notions they evolve in each case.

5. DISCUSSION

We have briefly considered composite signaling games where simple reinforcement learning allows for the evolution of basic pragmatic notions of *truth* and *probability*. In particular, we showed how it is possible for agents to coevolve terms that reliably indicate the truth value or probability of the coevolving descriptive base-game expressions. In each case, the pragmatic notion is characterized by the role it comes to play in the particular game.

---

[17]A cumulative success rate of better than 0.90 was used as the cutoff for success of the metagame. Note that here, as in the truth game, the base-game states that the metagame sender observes are biased. See note 10.
[18]A natural variant of the present model would be to make the metagame reinforcements proportional to the metagame receiver's actual winnings.

Such evolutionary models allow one to individuate between alternative notions of truth and probability. This way of characterizing alternative notions of truth and probability is pragmatic insofar as the notions are individuated by the role that they in fact play in facilitating successful action given whatever the aims of the agents may be. But there is another sense in which the particular variety of truth predicate that evolves in the model described above, when one does, serves as a pragmatic notion of truth *within the model*. The metagame in this model tracks whether the base game sender sends the customary signal. The customary signal routinely leads to successful action if and only if the base game evolves a reliable signaling system. So the metagame term that indicates that the customary signal was used in the base game represents a truth predicate if and only if customary signals in the base game in fact evolve to produce successful action. In this case, if a base-game signal gets assigned *true* then it will typically lead to a successful action. That is, here, *true* signals are the ones that work. Of course, while this is a feature of the particular notion of truth that evolves here, the more general point is that there are other pragmatic nations of truth characterized by other games where the predicates may come play quite different internal roles but still be recognized as truth-like in their behavior.

Finally, that the base language and the metalanguage *coevolve* in the composite games is significant. For an expression to be true or false, it must be meaningful. indeed, before there is a meaningful base language in the truth game, there is nothing that might be true or false. The metagame comes to track their truth as the base game expressions come to faithfully represent the world. Similarly, in the probability game, before there is a meaningful base language, there are no descriptions that might be probable or improbable. The significance of one's pragmatic notion of probability must depend on the significance of the descriptions to which it is applied.

Here, as in scientific practice, what is meant in taking a description to be true or in assigning a probability to a description coevolves with the significance of the description.[19] Further, on the present account, the meanings of one's basic descriptions and the pragmatic significance of associated notions of truth and probability become increasingly sharp together.

---

[19]See Barrett (2014) for a discussion of semantic drift in the context of empirical inquiry.

References

[1] Argiento, Raffaele, Robin Pemantle, Brian Skyrms and Stas Volkov (2009) "Learning to Signal: Analysis of a Micro-Level Reinforcement Model,"*Stochastic Processes and Their Applications* 119(2): 373–390.

[2] Barrett, J. A. (2015) "On the Evolution of Truth," forthcoming in *Erkenntis.* doi: 10.1007/s10670-015-9797-z. First online: 15 December 2015.

[3] Barrett, J. A. (2014) "On the Coevolution of Theory and Language and the Nature of Successful Inquiry," *Erkenntnis* 79(4): 821–834.

[4] Barrett, J. A. (2013) "On the Coevolution of Basic Arithmetic Language and Knowledge" *Erkenntnis* 78(5): 1025–1036

[5] Barrett, J. A. (2001) "Toward a Pragmatic Account of Scientific Knowledge," *PhilSci Archive.* uri: http://philsci-archive.pitt.edu/498/ accessed 20 August 2014.

[6] Barrett, J. A. and B. Skyrms (2015) "Self-Assembling Games," forthcoming in *The British Journal for the Philosophy of Science* first published online 13 September 2015. doi:10.1093/bjps/axv043.

[7] Barrett, J. A. and Kevin Zollman (2009) "The Role of Forgetting in the Evolution and Learning of Language," *Journal of Experimental and Theoretical Artificial Intelligence* 21(4): 293-309.

[8] Erev, I. and A. E. Roth (1998) "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria" *American Economic Review* 88: 848–81.

[9] Herrnstein, R. J. (1970) "On the Law of Effect," *Journal of the Experimental Analysis of Behavior* 13: 243–266.

[10] Hofbauer, Josef and Simon Huttegger (2008) "Feasibility of Communication in Binary Signaling Games," *Journal of Theoretical Biology* 254(4): 843–849.

[11] Huttegger, Simon, Brian Skyrms, Pierre Tarrès, and Elliott Wagner (2014) "Some Dynamics of Signaling Games," *Proceedings of the National Academy of Sciences* 111(S3): 10873–10880.

[12] Lewis, David (1969) *Convention.* Cambridge, MA: Harvard University Press.

[13] Roth, A. E. and I. Erev (1995) "Learning in Extensive Form Games: Experimental Data and Simple Dynamical Models in the Immediate Term," *Games and Economic Behavior* 8:164–212.

[14] Skyrms, Brian (2010) *Signals Evolution, Learning, & Information.* New York: Oxford University Press.

UC Irvine; Irvine, CA 92697, USA
E-mail address: `j.barrett@uci.edu`