COMMUNICATIVE BOTTLENECKS LEAD TO MAXIMAL INFORMATION TRANSFER

TRAVIS LACROIX

ABSTRACT. This paper presents new analytic and numerical analysis of signalling games that give rise to informational bottlenecks—that is to say, signalling games with more state/act pairs than available signals to communicate information about the world. I show via simulation that agents learning to coordinate tend to favour partitions of nature which provide maximal information transfer. This is true in spite of the fact that nothing from an initial analysis of the stability properties of the underlying signalling game suggests that this should be the case. As a first pass to explain this, I note that the underlying structure of our model favours maximal information transfer in regard to the simple combinatorial properties of the ways in which the agents might partition nature into kinds. However, I suggest that this does not perfectly capture the empirical results; thus, several open questions remain.

Keywords — signalling games, signals, signalling systems, information transfer, informational bottlenecks, reinforcement learning; emergent communication; sender-receiver games

1. INTRODUCTION

Descriptive language concerning *natural kinds* has been taken to evolve as a response to successful projections to more well-entrenched 'kind predicates' (Goodman, 1965). Furthermore, the degree to which such descriptive language (regarding terms that are supposed to *be* or *represent* kinds) is entrenched might have some bearing on 'general', as opposed to 'artificial', kinds. With this sort of analysis in mind, Barrett (2007) considers how a successful descriptive 'kind language' might evolve in the context of a simple sender-receiver game, called a *signalling game.*

The signalling game, due to Lewis (1969), shows how successful communication conventions might arise naturally by a process of repeated interactions. Evolutionary signalling games constitute a now-standard model for explaining and studying the emergence of communication in a wide range of social organisms—from humans and primates to bees and bacteria. This finds theoretical application in linguistics, evolutionary biology, and philosophy in research on language origins, in addition to practical application in computer science—especially machine learning and AI—in *emergent communication.* The basic signalling model has been extended in a number of ways in order to shed light on a variety of phenomena that are of philosophical interest, including meta-linguistic notions of truth and probability (Barrett, 2016,

DEPARTMENT OF LOGIC AND PHILOSOPHY OF SCIENCE, UNIVERSITY OF CALIFORNIA, IRVINE
MILA, (QUÉBEC AI INSTITUTE / INSTITUT QUÉBÉCOIS D'INTELLIGENCE ARTIFICIELLE)
E-mail address: tlacroix@uci.edu.

2017), self-assembly and modular composition (Barrett and Skyrms, 2017; Barrett et al., 2018; LaCroix, 2019b), prevarication (Skyrms and Barrett, 2018), etc.

Skyrms (2010a) highlights a variety of ways in which the simple signalling game framework might be extended. This may include varying the underlying dynamics, varying the parameters of the model, or imposing some sort of network structure on the game, among others. While there exist analytic results in certain cases, many such extensions are still being studied—often with the use of computer simulations.

I will be concerned here with a simple, but under-studied, extension of the signalling game where there are fewer signals than state/act pairs. This leads to an *informational bottleneck*, as it is impossible for the agents to learn a bijective mapping from states to signals and from signals to actions. That is to say, the agents in this sort of signalling game lack the expressive power to communicate full information about every possible state of the world that may obtain.

Of course, the players may learn rules for *syntactic combination* in order to perfectly partition nature. This is the situation discussed in Barrett (2006, 2007, 2009): signals are sent by multiple senders, which allows the receiver to obtain full information about the exact state of the world. In this case, the players must simultaneously invent the categories—i.e., partitions nature into distinct kinds— and code for representing the categories thus invented. Syntax, therefore, is helpful in *creating* 'new' signals from old ones in order to fully and naturally partition the world and represent those partitions via language. This is a step toward the sort of compositional richness that would be required for a fully *linguistic* communication system.[1]

The present paper offers several new analytic and experimental results regarding signalling situations where there are informational bottlenecks—i.e., fewer signals than state/act pairs. Rather than modifying the game by adding players, as Barrett (2006, 2007, 2009) does, I examine the *types* of partitions that in fact arise under numerical simulation.[2] Specifically, I demonstrate how prior analyses of evolutionary stability do not give a complete picture of the expected outcome of the dynamic in question—in this case, simple reinforcement learning. Perhaps surprisingly, there is a natural movement toward optimal information transfer. Thus, I examine why we should expect communication to naturally favour maximal information transfer in addition to rates of successful communication, in spite of the fact that extant analytic results (and indeed, the learning dynamic itself) only takes account of the latter.

Section 2 provides some background on the signalling-game framework. I offer several formal definitions that will be of some use later in the paper. Section 3 looks more closely at the signalling context that I will focus on here and surveys some extant analytic results, primarily due to Donaldson et al. (2007), to which I will appeal. Section 4 offers new analytic results, in addition to simulation results, to

---

[1]Though this simple syntactic game is not itself compositional, see the discussion in Franke (2016); Steinert-Threlkeld (2016); LaCroix (2019a).

[2]In a related paper, Barrett and LaCroix (2019) use these results to explain how the structural properties of a language come to reflect the world in which the language evolved. This shows how something like a principle of indifference (in a Bayesian sense) might arise naturally in an evolutionary context. This is discussed in further detail, in relation to the current analysis, in Section 5.

demonstrate the connection between communicative success and information transfer in a signalling-game context with fewer signals than state/act pairs. Section 5 concludes.

2. SIGNALLING GAMES

The Lewis signalling game has two players, whom we call the 'sender' and the 'receiver'. In the simplest signalling game, there are two possible states of the world, denoted by 's_0' and 's_1'; two possible signals or messages that the sender might select, denoted by 'm_0' and 'm_1'; and two possible actions that the receiver might perform, denoted by 'a_0' and 'a_1'. We refer to this as a 2×2 signalling game, corresponding to the number of state/act pairs by the number of signals. Here, each action coincides with a particular state of the world.

On each play of the game, nature picks a state of the world at random.[3] The sender observes the state of the world directly and randomly selects a message to send to the receiver. The receiver knows which message was sent, but does not know which state of the world obtains. She then chooses an action at random. If the receiver's action was appropriate for the state, then the play is successful; otherwise it is a failure.

The players have evolved or learned an efficient language when they perform better than chance on coordinating signals to state/act pairs. A maximally efficient signalling convention is referred to as a *signalling system* (Lewis, 1969). In this case, there is a map from unique states to unique signals to appropriate actions. The 2×2 signalling game has two possible signalling systems, shown in Figure 1. In

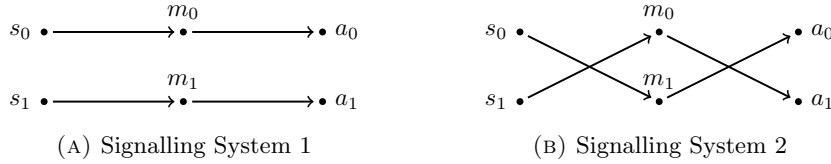(A) Signalling System 1   (B) Signalling System 2

FIGURE 1. The two signalling systems of the 2×2 signalling game

either case, at a signalling system for the 2×2 signalling game, the sender and receiver have a communicative success rate of 1, and each signal carries exactly 1 bit of information.

The way in which the sender and receiver achieve such a communication convention is going to depend largely upon the dynamic under consideration.[4] I will use a simple learning dynamic, called *Herrnstein reinforcement learning*. On this dynamic, the probability of an actor's choosing a particular action is proportional to that action's accumulated rewards.[5] When the players succeed, they earn some

---

[3]For now, I will assume that all states are equiprobable, so nature is *unbiased*, but this assumption can be relaxed.

[4]See Huttegger et al. (2010) for an overview of a number of different dynamics.

[5]This simple reinforcement learning is based upon the 'matching law' proposed by Herrnstein (1970), which is itself a formalisation of the 'law of effect', due to Thorndike (1905, 1911, 1927). This learning rule is empirically tested in Roth and Erev (1995); Erev and Roth (1998). On the real-world effectiveness of simple learning, Schultz et al. (1997) show that dopamine neurons in certain areas of primate brains seem to enact a reasonably similar learning procedure. See also Schultz (2004) and Glimcher (2011).

payoff and thus reinforce their behaviour. When they fail, they may receive nothing, or they may receive some negative payoff (punishment).

We can understand this intuitively as a type of urn-learning procedure. For the 2×2 signalling game, we now suppose that the sender comes equipped with two urns—one labelled 's_0', and the other labelled 's_1'. Each of the sender's urns contains two balls—labelled 'm_0' and 'm_1', respectively. Similarly, the receiver comes equipped with two urns—one labelled 'm_0', and the other labelled 'm_1'—and each of her urns contains two balls—labelled 'a_0' and 'a_1', respectively. See Figure 2.

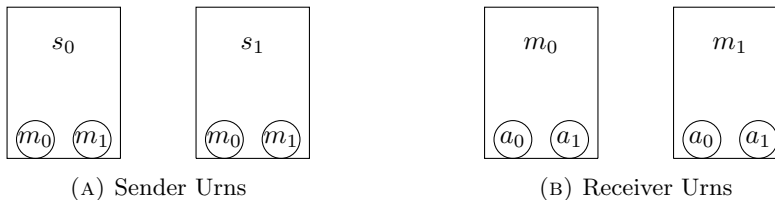(A) Sender Urns   (B) Receiver Urns

FIGURE 2. Simple reinforcement learning model

Now, on each play of the game, nature picks a state of the world at random. The sender then chooses a ball at random from the urn corresponding to the current state of the world and sends the message chosen to the receiver. The receiver chooses a ball at random from the urn corresponding to the message that she received. If the action chosen matches the state of the world, then both the sender and receiver return their ball to the urn from which it was chosen and additionally reinforce their behaviour by adding another ball of the same type to that urn, thus shifting the probabilities of choosing a particular signal or action over future plays. If the action does not match the state of the world, then the sender and receiver both simply return the ball to its urn.[6] The game is then repeated.

This reinforcement procedure makes it more likely that a message [act] that led to a success will be chosen when a particular state [message] occurs. Thus, signals come to carry information (Skyrms, 2010a,b). Though this model is relatively simple, it turns out to be extremely effective for learning how to communicate.[7]

The information that a signal carries might be understood as information *about* the state.[8] When signals are completely informative, the receiver has complete information about the state of the world, and so can act as though she had observed the state directly. Skyrms (2010a,b) shows how we can specify the quantity of information in a given signal using Kullback-Leibler divergence. This depends upon a comparison between the (conditional) probability that the agents are in a

---

[6]Punishment for miscoordination can be included in this model. This simply consists in discarding the ball when the agents fail to coordinate.

[7]In the 2×2 signalling game, when nature is not too biased, the sender and receiver converge toward one or the other signalling system with probability 1 under this sort of dynamic (Argiento et al., 2009). Therefore, in the limit, the players learn to coordinate perfectly. Furthermore, after only 300 time-steps, the communicative success rate of the sender and receiver is approximately 0.9, on average (Skyrms, 2010a). Similar results hold for a variety of other dynamics, such as the replicator dynamic. See Huttegger (2007a,c) for an overview.

[8]The signal may also be understood as being about actions; see Huttegger (2007b); Zollman (2011) on separating indicatives and imperatives.

particular state *given that* a signal was sent, and the probability that the agents are in that state simpliciter.[9]

Thus, we can define the *quantity* of information carried by a particular signal, m_j (i.e., about a particular state, s_i) as

$$H(m_j) = \log \frac{p(s_i|m_j)}{p(s_i)}.$$

This quantity at a signalling system results in m_j carrying 1 bit of information in the 2×2 game.[10]

Skyrms (2010a,b) points out that signals may carry some information about different states. Thus, in order to get a real sense of the amount of information that a particular signal carries, we can take a weighted sum of the probabilities of being in any of the particular states conditional upon the particular signal. Thus, we obtain the following measure of the quantity of information that is carried by a particular signal, m_j:

$$I(m_j) = \sum_{i=1}^{|S|} p(s_i|m_j) \cdot \log \left(\frac{p(s_i|m_j)}{p(s_i)} \right).$$

Note that this definition requires a slightly different interpretation than the original notion of Shannon entropy, since we are examining the quantity of information given by a single message.[11] This is a notion of *semantic information* (Dretske, 1981). In this context, the relative entropy of a particular signal can be understood as a measure of the information with respect to *additional bits gained* by moving from a prior to a posterior distribution, in a Bayesian sense.

We might also note the total entropy of a signalling strategy. This is just the sum of the information contained in all of the individual messages, weighted by the probability that a message occurs. In the case of a signalling system in a symmetric $n \times n$ signalling game, the probability that a message occurs will just equal the probability that a given state (the one represented by that message) will occur.

For our purposes, it will be convenient to have a formal definition of a signalling game, thus.[12]

> **Definition 2.1:** *Signalling Game*
> Let $\Delta(X)$ be a set of probability distributions over a finite set X. A *Signalling Game* is a tuple
>
> $$\Sigma = \langle S, M, A, \sigma, \rho, u, P \rangle$$
>
> where $S = \{s_1, \ldots, s_k\}$ is a set of *states*, $M = \{m_1, \ldots, m_l\}$ is a set of *messages*, $A = \{a_1, \ldots, a_n\}$ is a set of *acts*, with S, M, and A nonempty. $\sigma : S \to \Delta(M)$, is a function which defines a *sender*, $\rho : M \to \Delta(A)$ defines

---

[9]See Godfrey-Smith (2011); Birch (2014); Shea et al. (2018) for a critical discussion of Skyrms' notion of informational content in the signalling game.

[10]That is to say, there is a reduction of uncertainty from two possibilities to one. At a signalling system in a 4×4 game, signals carry two bits of information (and so on). One bit of information, in this example, corresponds to the logarithm's base being 2. The units of this quantity can be specified similarly as *nats* or *harts* if we change the base to e or 10, respectively.

[11]See, Shannon (1948) and Shannon and Weaver (1949).

[12]See Huttegger (2007a) and Steinert-Threlkeld (2016).

a *receiver*, $u : S \times A \to \mathbb{R}$ defines a *utility function*, and $P \in \Delta(S)$ gives a probability distribution over states in S. Finally, σ and ρ have a common *payoff*, given by

$$\pi(\sigma, \rho) = \sum_{s \in S} P(s) \sum_{a \in A} u(s, a) \cdot \left(\sum_{m \in M} \sigma(s)(m) \cdot \rho(m)(a) \right)$$

The payoff $\pi(\sigma, \rho)$ for a particular combination of sender and receiver strategies is an expectation of the utilities of state/act pairs (given by $u(s, a)$) weighted by the relative probability of a particular state, given by $P(S)$. This is the *communicative success rate* of σ and ρ. With this definition in mind, the signalling systems of a signalling game can be defined formally as follows.

> **Definition 2.2:** *Signalling Systems*
> A signalling system in a signalling game is a pair (σ, ρ) of a sender and receiver that maximises $\pi(\rho, \sigma)$.

This definition of a signalling game is more general than, e.g., that of Lewis (1969). I will be interested here in more narrow contexts of coordination problems. In particular, I will focus on signalling contexts where there are fewer messages than state/act pairs, where nature is unbiased, and where the utility function pays 1 just in case the state and the act match and zero otherwise. Thus, the $n \times m$ signalling game with unbiased nature is defined as follows.

> **Definition 2.3:** $n \times m$ *Signalling Game With Unbiased Nature*
> The $n \times m$ *Signalling Game* is the signalling game with $|S| = |A| = n$, $|M| = m$, $u(s_i, a_j) = \delta_{ij}$, and $P(s) = 1/n$ for every $s \in S$. Further, δ_{ij} is the *Kronecker delta*, defined as
>
> $$\delta_{ij} = \left\{ \begin{array}{ll} 1 & \text{if } i = j \\ 0 & \text{else} \end{array} \right. .$$

Note, that the symmetric $n \times n$ signalling game is just a special case of the $n \times m$ signalling game with $|S| = |M| = |A| = n$. If $n < m$ we have a situation wherein synonyms might arise (Skyrms, 2010a; Hu et al., 2011). If $n > m$ we get a situation where it is impossible for the sender and receiver to communicate perfectly. This gives rise to an informational bottleneck. I will be concerned here with solely the situation that is described in Definition 2.3, with $n > m$. For ease of exposition, I will refer to this as an *impoverished* signalling game, since the communicative resources of the sender are impoversihed with respect to the complexity of the world.

Donaldson et al. (2007) offer an analytic analysis of this sort of situation; however, they use a more general utility function, $u(s_i, a_j)$, than the Kronecker delta, so that there may be, for example, a general-purpose action that is more appropriate than a specific-purpose action when the state of the world is uncertain. Advancing the earlier work of Wärneryd (1993) and Trapa and Nowak (2000), they show that it is possible to define an evolutionarily stable *set* of strategies in such a context. This is particularly important for situations modelling animal communication since "real predator alarm call systems tend to employ only a few signals to distinguish

between predators, with many types lumped together" (Donaldson et al., 2007, 231). In such a case it is advantageous for individuals represent a pool of states by a single signal when those states are functionally similar. Based upon the analysis of Donaldson et al. (2007), I present some useful properties of the impoverished signalling game in Section 3.

3. State Partitions and Informational Bottlenecks

This section summarises and analyses some previous results for informational bottlenecks that are pertinent to the discussion herein. In an evolutionary context, we are effectively concerned with a two-population model, and thus a role-asymmetric game. As such, a sender strategy in an evolutionarily stable state (ESS) must be *uniquely* optimal against the receiver strategy, and the receiver strategy must, similarly, be uniquely optimal against the sender strategy (Selten, 1980).

Donaldson et al. (2007) show that the following four conditions are necessary for a signalling system to be a strict Nash equilibrium, and therefore an evolutionarily stable strategy.

Property 3.1: The signalling strategy, σ, must be binary; that is, each state gives rise to exactly one signal.[13]

Property 3.2: The receiving strategy, ρ, must be binary; that is, each signal results in exactly one action.[14]

Property 3.3: The signalling strategy, σ, must be onto; that is, every signal must be used.[15]

Property 3.4: The receiving strategy, ρ, must be one-to-one; that is, no two signals may give rise to the same action.[16]

Donaldson et al. (2007) note that these four properties limit the *type* of multiplicity that is allowable in a signalling game. One state leading to multiple signals, and one signal leading to multiple actions are both ruled out by a theorem of Selten (1980), from which Properties 3.1 and 3.2 follow. Namely, an evolutionarily stable strategy can only have one possible response for a particular circumstance. Property 3.4 rules out the possibility that multiple signals lead to the same action. However, the possibility of multiple states leading to the same signal is not ruled out. Donaldson et al. (2007) point out that if this is the case, then the signal comes to 'mean' that one of several possible situations has occurred. (They further point out that if payoffs are asymmetric, it is still possible to determine the unique action with maximal payoff.) Though Properties 3.1–3.4 are individually necessary for evolutionary stability, they are not jointly sufficient.

I further introduce the following definitions, assuming Properties 3.1–3.4 hold, to obtain a formal description of a state partition, where there are fewer signals than state/act pairs, and a best response to such a partition.

---

[13]See Donaldson et al. (2007). This follows from a proof due to Selten (1980).
[14]See Donaldson et al. (2007). This follows from a proof due to Selten (1980).
[15]Proof is given by Donaldson et al. (2007).
[16]Proof is given by Donaldson et al. (2007).

Definition 3.1: *State Partition*
The *partition* of states $\tau(m)$ associated with a signal m is the set of states mapping to that signal under σ: $\tau(m) = \{s : p(m|s) = 1\}$.[17]

Definition 3.2: *Best Response to a State Partition*
A best response to a partition of states is an action which maximises the expected payoff for all members of the partition:

$$\mathrm{BR}(\tau) = \arg\max_{a \in A} \sum_{s \in \tau} p(s)u(s,a).$$

If there is a unique such action, it is termed the strict best response to the partition.

Given these definitions, two further properties hold:

Property 3.5: Every partition must have a strict best response, and the signal corresponding to that partition must map to it.[18]

Property 3.6: For each member of a partition of states, the strict best response for that partition must be a better response than the strict best response of any other partition. That is, for all $s \in \tau(m_i), m_j \neq m_i$:[19]

$$\pi(\sigma, \mathrm{BR}(\tau(m_i))) > \pi(\sigma, \mathrm{BR}(\tau(m_j))).$$

Thus, Donaldson et al. (2007) prove the following general theorem for evolutionary stability:

Theorem 3.1: A strategy that is determined by σ, ρ is evolutionarily stable if and only if the properties 3.1–3.6 listed above hold.[20]

In the case where there are fewer signals than state/act pairs, as in the impoverished signalling game, multiple states necessarily map to a single signal. While this situation is consistent with the above properties, when payoffs are symmetric it is impossible for this situation to be evolutionarily stable, since there cannot be a unique (strict) best response to that partition—i.e., Property 3.6 is violated. However, an evolutionarily stable *set*—namely, a partition that groups together a set of acts and has a well-defined probability distribution over the acts—satisfies Properties 3.1 – 3.6 and thus is evolutionarily stable. Figure 3 shows a particular example where $\tau(m_0) = \{s_0, s_1\}$ and $\tau(m_1) = \{s_2\}$. No single evolutionarily stable strategy exists for this system. However, the set of strategies given by

$$\{(\sigma, \rho(x)) : x \in [0,1]\}$$

is itself evolutionarily stable, since it constitutes an evolutionarily stable set.

---

[17]Donaldson et al. (2007) refer to this as a 'pool', in order to draw a connection to pooling equilibria. I will refer to τ as a partition. Thus, 'pool' and 'partition' refer to the same thing.
[18]Proof is given by Donaldson et al. (2007).
[19]Proof is given by Donaldson et al. (2007).
[20]Proof is given by Donaldson et al. (2007).

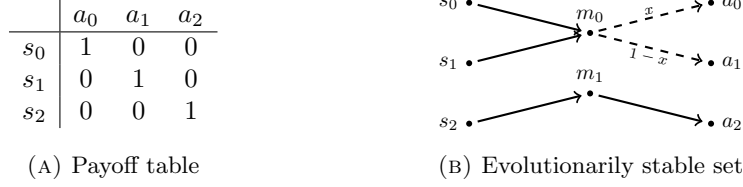|	a_0	a_1	a_2
s_0	1	0	0
s_1	0	1	0
s_2	0	0	1

(A) Payoff table

(B) Evolutionarily stable set

FIGURE 3. Example of an evolutionarily stable set of strategies

4. COMMUNICATIVE SUCCESS AND INFORMATION TRANSFER

In this section, I present novel analytic results for the impoverished signalling game. However, the analytic results (Section 4.1) highlight that all stable partitions have equivalent expected payoff. As such, we might suppose that each of these possibilities is equally likely. When we move to the simulation results of Section 4.2, however, we see that this is not the case. Thus, the main results of this paper highlight that the analytic treatment, e.g., of Donaldson et al. (2007) is insufficient for an explanation of what is happening on these models. In particular, the players tend to favour partitions that offer maximal information transfer. This phenomenon is not explained by previous analytic results. I discuss a possible combinatorial explanation of what is driving this result; however, I show that this too is not sufficient to explain these outcomes.

4.1. **Analytic Results.** The first proposition that we will require is given as follows.

> **Proposition 4.1:** The maximal communicative success rate for the impoverished signalling game, as defined in Definition 2.3, is
> $$\max \pi(\sigma, \rho) = \frac{|M|}{|S|},$$
> where $|M|$ is the cardinality of the set of messages, and $|S|$ is the cardinality of the set of states.
>
> *Proof.* See Appendix A for details. □

Note that for the symmetric $n \times n$ signalling game, this reduces to a maximum expected payoff of 1 ($|M| = |S|$), as would be expected. Two corollaries follow immediately from this.

> **Corollary 4.1:** In a signalling game, as defined in Definition 2.3, a strategy wherein the sender does not take advantage of all of the signals available to her has a lower communicative success rate than a strategy which does take advantage of all the signals available to her.
>
> *Proof.* See Appendix A for details. □

Given this fact, it should not matter how many states are in any partition as long as there are the same number of partitions as there are signals. (This also implies that every partition contains at least one state.) Further:

Corollary 4.2: All probability distributions of strategies that take advantage of all available messages have equivalent (maximal) communicative success.

> *Proof.* Since I assumed nothing about the actual probability distributions in the proof of Proposition 4.1, this follows immediately. □

Thus, it does not matter what the receiver's probabilities are as long as she partitions the acts in the same way that the sender partitions the states.

What I have just shown is that, under the assumption that nature is unbiased (all states are equiprobable), and the utility function is given by $u(i,j) = \delta_{i,j}$—the Kroenecker delta—as long as the sender takes advantage of all possible signals at her disposal, the sender and receiver both receive maximal payoff, $\max \pi(\sigma, \rho) = \frac{|M|}{|S|}$, independently of *how* the states are partitioned and independently of the receiver's probability distribution over acts given signals.

This gives rise to the following question: which partitions and which receiver strategies (in response to those partitions) might we expect to arise for this type of signalling game? The analytic results offer no indication of an answer to this question—as long as a partition is stable, it will constitute a signalling system.

Based on the preceding analysis alone, we might expect that, for a signalling game with more state/act pairs than messages, every possible partition should show up with uniform probability. However, the reasoning for such an intuition is flawed, since not every partition is equally probable.

Recalling that $\tau(m_i) = \{s_{i,0}, s_{i,1}, \ldots, s_{i,j}\}$ denotes a set of states that map to a message, m_i, the cardinality of $\tau(m_i)$ tells us how many states are 'contained' in that partition. Thus, let us introduce the following notation for denoting a full partition of nature:

$$\tau(M) = \langle |\tau(m_0)|, |\tau(m_1)|, \ldots, |\tau(m_i)| \rangle$$

For example, the unbiased partition for the 4×2 signalling game is denoted by $\langle 2, 2 \rangle$, which can be read in the following way: the first partition, corresponding to m_0, contains 2 states, and the second partition, corresponding to m_1, contains 2 states.

Taking combinatorial properties into account, we note that there is only one way at arriving at the unbiased partition, $\langle 2, 2 \rangle$, but there are two ways of arriving at each of the biased partitions, $\langle 3, 1 \rangle$ versus $\langle 1, 3 \rangle$ and $\langle 4, 0 \rangle$ versus $\langle 0, 4 \rangle$. Since the latter pair do not take advantage of all of the signals available, we can assume that they will not occur. (Note that these outcomes are unstable, as per Section 3.) Thus, there are three possible outcomes that we should expect, and I consider these to be distinct partitions. Thus, the answer to the question suggested above would indicate that we might expect each of these partitions to occur with frequency 1/3, since they all have the same communicative success rate (under the assumption that the receiver is best responding to the partition).

However, this is not the case: this does not account for the fact that each 'distinct' partition, in the above sense, has a variety of ways that it might be arrived at, and these are not equivalent. In the case of 4 states, we have $\binom{4}{2} = 6$ ways of arriving at the unbiased partition and $\binom{4}{1} = \binom{4}{3} = 4$ ways of arriving at each

of the biased partitions. Therefore, there are 14 possible combinations to consider, and their relative frequencies are $3/7$, $2/7$ and $2/7$, respectively.[21]

Thus, depending on how we cash out 'distinct', we should either expect a distribution that looks like a binomial coefficient—in case I call $\langle 2, 2\rangle$, $\langle 3, 1\rangle$ and $\langle 1, 3\rangle$ 'distinct'. This way of cashing out 'distinct' does not account for *which* states are in a given partition, but rather which *types* of partitions occur. This is a fairly coarse-grained analysis. However, if we examine a more fine grained analysis, we will expect a uniform distribution across the various *ways* of getting at each partition—i.e., taking account of which states are in a given partition, and calling that distinct.

So, the reasoning is that we should expect on simulation to see a distribution of partitions that approximates a binomial coefficient. Surprisingly, this is not the case. What we find is that the partitions tend to be biased toward maximal information transfer—more weight is put on those partitions close to the unbiased partition. I present the data from these simulation results below and discuss why the the combinatorial measure alone cannot be used to explain the partitions that evolve under simple reinforcement.

4.2. **Simulation Results.** In this section, I present simulation results which show that neither an analysis of the stability properties of the impoverished signalling game, nor the combinatorial argument given above can serve to explain the observed distribution of partitions under simple reinforcement learning.

4.2.1. *Unbiased Nature.* We will start with a 10×2 signalling game with unbiased nature. Thus, there are 10 states, each equiprobable, and 10 correspondent actions, but only 2 signals with which the sender may represent these states.[22] The chance payoff for this game is $1/10 = 0.1$, and the maximal payoff—occurring at any possible stable partition and correspondent receiver strategy, as per the analytic results above—is $1/5 = 0.2$. After 10^6 plays per run, the expected payoff is 0.1999, on average (1000 runs). Every run (1.000) achieves an expected payoff greater than 0.1990. Furthermore, learning is fast. This is demonstrated by the fact that the cumulative success rate is 0.1995 after 10^6 plays per run, on average (1000 runs).[23] Many runs (0.875) achieve a cumulative success rate of at least 0.1990.

As was suggested above by the reasoning that appealed to the combinatorial measure, the unbiased $\langle 5, 5\rangle$ partition is indeed the most common of the 9 stable partitions for this game: almost $1/3$ of the time (0.310), the sender and receiver perfectly partition nature for maximal information transfer. Most of the time (0.792), the sender and receiver partition nature *near*-perfectly—that is, a $\langle 5, 5\rangle$, $\langle 6, 4\rangle$, or $\langle 4, 6\rangle$ partition. Rarely (0.031), the sender and receiver fail to evolve a clear partition of nature.

---

[21]Again, I do not include in our frequency expectation the possible outcome for $\binom{4}{0} = \binom{4}{4} = 1$. This would have us calculate our expected frequencies out of 16 possibilities rather than 14; however, these outcomes correspond to unstable strategies, and so they will be selected against in the long-term.

[22]In our urn-learning metaphor, the sender has ten urns, each starting with two balls, and the receiver has two urns each starting with 10 balls.

[23]The cumulative success rate is a measure of success that takes account of the history of the game. It is calculated by dividing the number of plays that led to a success by the total number of plays in that run. When the players are successful, early failures are washed out as the number of plays increases.

Figure 4 shows the distribution of runs for the impoverished signalling game with 10 state/act pairs and 2 signals under simple reinforcement. The entropy (mean information per signal) is represented in this figure simply to illustrate how the bias of the partitions correspond to information transfer—namely, the less biased the partition, the more informative it is. Note that the combinatorial measure does
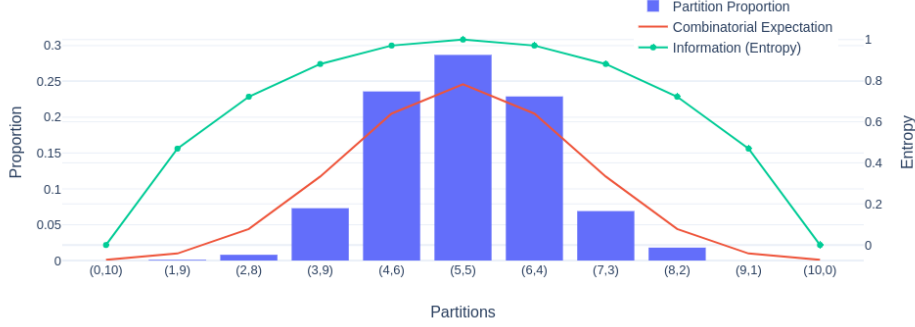

FIGURE 4. Partitioning 10 unbiased states with 2 signals. Comparison of experimental results with combinatorial expectation and average information transfer

not perfectly track the actual distribution of partitions on simulation. One might attribute this to mere noise; however, the likelihood of more biased partitions drops off even faster than the number of ways to obtain these partitions.

We see the same phenomena in the impoverished signalling game with 9 state/act pairs and 3 signals on simulation under simple reinforcement learning. In this case, the chance payoff is $1/9 \approx 0.1111$, and the maximal payoff—again, occurring at any possible stable partition and correspondent receiver strategy—is $1/3 \approx 0.3333$. After 10^6 plays per run, the expected payoff is 0.3332, on average (1000 runs). Almost every (0.926) run achieves a near-perfect expected payoff greater than 0.3330, and every run (1.000) achieves an expected payoff greater than 0.3300. Again, learning is fast: the cumulative success rate is 0.3323 after 10^6 plays per run, on average, and almost every run (0.990) achieves a cumulative success rate of at least 0.330.

Again, the unbiased partition, $\langle 3, 3, 3 \rangle$, is the most common of the 28 stable partitions. About 1/6 of the time (0.160), the sender and receiver perfectly partition nature for maximal information transfer, and more than 2/3 of the time (0.684), the sender and receiver partition nature near-perfectly—that is, a $\langle 3, 3, 3 \rangle$ partition or one of the six permutations of the $\langle 4, 3, 2 \rangle$ partition. Rarely (0.031), the sender and receiver fail to evolve a clear partition of nature—usually, this is because the receiver has learned early on to never choose some act. In this case, the sender usually mixes the correspondent state over two signals. The urn for this state (almost) never changes, since the receiver (almost) never chooses that action.[24]

Figure 5 shows the distribution of runs for the impoverished signalling game with 9 state/act pairs and 3 signals under simple reinforcement. The fact that the

[24]Of course, since there is no punishment, every conditional action for the sender and receiver has *some* weight; however, in these cases the weight for these acts is less than 1 in 2500.)
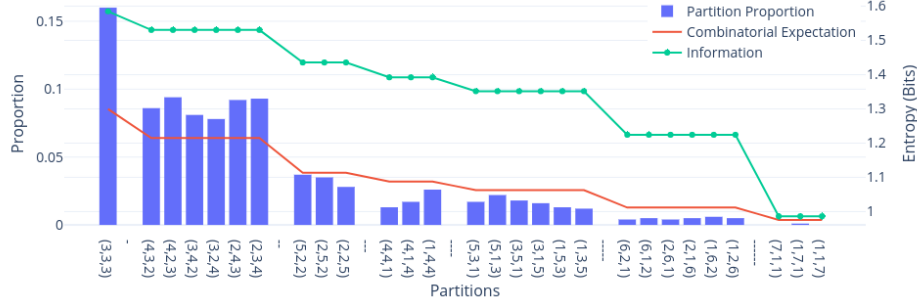
FIGURE 5. Partitioning 9 unbiased states with 3 signals. Comparison of experimental results with combinatorial expectation and average information transfer

combinatorial measure does not perfectly track the actual distribution of partitions on simulation is even more pronounced here: the observed frequency of the $\langle 3, 3, 3 \rangle$ partition is twice that of the expected frequency based on the combinatorial argument. This suggests that these discrepancies cannot be accounted for by mere noise. Again, the likelihood of more biased partitions drops off even faster than the number of ways to obtain these partitions: the observed frequency of the $\langle 6, 2, 1 \rangle$ partition, for example, is $1/4$ of the expected frequency.

As such, it appears that the impoverished signalling game tracks unbiased—or close to unbiased—partitions significantly better than the combinatorial measure over the possible ways of obtaining each distinct type of partition. When we introduce some bias into nature, we see that the combinatorial measure alone cannot be used to explain the observed frequencies.

4.2.2. *Biased Nature.* Let us consider again the impoverished signalling game with 10 state/act pairs and 2 signals. However, let us now suppose that nature is biased so that the probability distribution over the set of states is given by

$$\Delta(S) = \begin{bmatrix} \frac{1}{2} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} & \frac{1}{18} \end{bmatrix}$$

for states s_0 through s_9, respectively. Again, the chance payoff is $1/10 = 0.10$. When signals are used optimally—i.e., a $\langle 1, 9 \rangle$ partition, corresponding to the biased state and every other state—the maximal expected payoff is $5/9 \approx 0.5556$. On simulation, under simple reinforcement, after 10^6 plays per run, the expected payoff is 0.5502, on average (1000 runs). Many runs (0.8989) achieve an expected payoff of at least 0.550 by 10^6 plays. As with the unbiased-nature simulations, learning is fast: the cumulative success rate is 0.5496 after 10^6 plays, on average (1000 runs), and many runs (0.887) achieve a cumulative success greater than 0.550.

Most of the time (0.897) the players learn to coordinate upon a partition nature. Most often (0.570), we see a $\langle 9, 1 \rangle$ or a $\langle 1, 9 \rangle$ partition (with approximately equal frequency). Here, the sender partitions the biased state into one signal, and the rest of the states into the other signal. About $1/4$ of the time (0.281), the sender pools a second state into the signal containing the biased state for an $\langle 8, 2 \rangle$ or

$\langle 2, 8 \rangle$ partition (again, with approximately equal frequency). Sometimes (0.082), the sender pools two extra states into the signal containing the biased state for a $\langle 3, 7 \rangle$ or $\langle 7, 3 \rangle$ partition. Rarely (0.008), the sender pools three extra states into the signal containing the biased state for a $\langle 4, 6 \rangle$ or $\langle 6, 4 \rangle$ partition. On simulation, we never see a $\langle 5, 5 \rangle$ partition.

However, in *every* case, when the sender pools an extra state into the signal that contains the biased state, the receiver learns to never choose the action corresponding to that signal. That is, the sender always partitions *one* action—the action corresponding to the biased state—via the signal that contains the biased state in its partition. Thus, when the partition is, e.g., $\langle 2, 8 \rangle$, the signal that pools two states carries disjunctive information about those states, and is sent with (unconditional) probability 10/18; however, the receiver interprets the signal as carrying complete information about the act—namely, the act corresponding to the biased state.

More subtle pooling can occur when nature is biased in this way. Sometimes (0.083), the receiver learns to ignore the signal, and only do the action corresponding to the biased state. (For one signal, she puts full weight on the act corresponding to the biased state; for the other signal, she puts weight between $(0.95, 1.00)$ on the same state, with weight between $(0.00, 0.05)$ distributed over the remaining actions for that signal.) Here the players receive an average payoff of 0.4999.

The rest of the time (0.020), we get a graded result between these two cases. The sender puts full weight on the action corresponding to the biased state for one signal, and she puts some weight between $(0.05, 0.95)$ on the same act for the other signal. This gives an average payoff of 0.5222.

We see similar behaviour when we examine the impoverished signalling game with 9 state/act pairs and 3 signals. Suppose that nature is biased so that the probability distribution over the set of states is given by

$$\Delta(S) = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{21} & \frac{1}{21} & \frac{1}{21} & \frac{1}{21} & \frac{1}{21} & \frac{1}{21} & \frac{1}{21} \end{bmatrix}$$

for states s_0 through s_8, respectively. The chance payoff, again, is $1/9 \approx 0.1111$. When signals are used optimally—i.e., a $\langle 1, 1, 7 \rangle$ partition—the maximal expected payoff is $15/21 \approx 0.7142$. On simulation, under simple reinforcement, after 10^6 plays per run, the expected payoff is 0.6970, on average (1000 runs), and many runs (0.794) achieve an expected payoff of at least 0.71 by 10^6 plays. The cumulative success rate is 0.6952 after 10^6 plays, on average (1000 runs), with about 3/4 (0.742) of the runs achieving a cumulative success rate of at least 0.71.

As with the 10 case with biased nature, qualitatively, three different things happen. Most of the time (0.772), the sender and receiver optimally partition the states/acts with the signals for a maximal possible payoff. Of these, the most common result (0.347) is that the sender partitions one biased state into one signal, the other biased state into the other signal, and the remaining seven states into the third signal—a $\langle 7, 1, 1 \rangle$ partition (or its permutations, with each permutation occurring with approximately equal frequency), and the receiver responds accordingly. Next most often (0.325), the sender pools an extra state with one of the signals that contains one of the biased states for a $\langle 6, 2, 1 \rangle$ partition. The rest

of the time (0.194), we see all the remaining partitions, noting that the unbiased $\langle 3, 3, 3 \rangle$ partition occurs least frequently.[25]

The most important thing to note is that when the sender pools another state (or states) into the signal containing one of the biased states, the receiver learns to ignore the action associated with the extra state, and puts all weight for that signal onto the act associated with the biased state; thus, as with the 10×2 case, even though the signal may carry disjunctive information about the states, the receiver interprets it as carrying full information about a single act.

Sometimes (0.034), the sender pools both of the biased actions into one signal, and the receiver mixes over the appropriate actions, given that signal. The receiver perfectly partitions the remaining seven states into the other two signals, and the receiver mixes over the appropriate actions, conditional upon the signal received. This strategy has an expected payoff of $9/21 \approx 0.4286$.

Again, the rest of the time, we get a graded result between these two cases. Here the sender puts some weight on the same (biased) state over two signals, and the receiver puts some weight on the same action (corresponding to that biased state) from two signals. Thus, the sender mixes signals for one biased state and perfectly partitions the other biased state. When the receiver puts a lot of weight on the same act for two different signals, she generally ignores all the other actions. This strategy has an expected payoff of $2/3 = 0.6667$. This happens under $1/5$ of the time (0.194).

4.2.3. *Different Dynamics.* We might wonder whether there is something special about simple reinforcement learning that leads to the results we have seen thus far. However, in examining variations of the learning dynamic, we must be sensitive to the parameters of this game. For example, under a dynamic like Win-Stay/Lose-Shift, no set of strategies will be stable. Since there is an significant informational bottleneck in the 10 impoverished signalling game, at a signalling system the best the players can do is to lose half the time. Thus, they will constantly shift their strategies.[26] However, they may spend some time in one region of the distribution space, depending upon how the parameters of the model are actually implemented.

Here, we examine a variation of the simple reinforcement learning model to include punishment for miscoordination. However, the model is extremely sensitive to the parameters. We have already seen, in the simple reinforcement learning model, that the sender and receiver often co-evolve their strategies; however, the receiver sometimes learns more quickly than the sender, and this results in the receiver ignoring some (or, when nature is quite biased, most) of the signals that the sender sends. Therefore, if punishment is too severe, then the receiver will quickly learn to glom onto a particular action for one or the other signal. This results in a $\langle 1, 9 \rangle$ or $\langle 9, 1 \rangle$ partition on the receiver's end. However, the sender has no opportunity to finesse a partition to which the receiver will be responsive. If the receiver learns too quickly—as happens with severe punishment—she never

---

[25]More specifically, this includes the $\langle 5, 3, 1 \rangle$ partition (occurring 0.073); the $\langle 4, 4, 1 \rangle$ partition (occurring 0.019); the $\langle 5, 2, 2 \rangle$ partition (occurring 0.066); the $\langle 4, 3, 2 \rangle$ partition (occurring 0.030); and, rarely, the $\langle 3, 3, 3 \rangle$ partition (occurring 0.005). Again, in each case we see all permutations occurring with approximately equal frequency.

[26]See discussion of Win-Stay/Lose-Shift in Barrett and Zollman (2009); Huttegger and Zollman (2011), and the related dynamic Win-Stay/Lose-Randomise in Barrett and Zollman (2009); Barrett et al. (2017).

performs certain actions, so the sender is continually punished, at equal rates for each of the signals, when the states corresponding to those unused actions are chosen by nature. Thus, on average, her strategy remains unchanged for those particular states.

We examine a 10×2 signalling game where nature is unbiased, and where the dynamic includes a positive payoff $[+2]$ for coordination, and a negative payoff, or punishment, $[-0.1]$ for miscoordination. The success rate is as before—0.10 at chance, and 0.20 at a signalling system. Taking account of reward and punishment, the expected payoff is 0.11 at chance and 0.32 at a signalling system. Though the formula for calculating the maximum expected payoff at a signalling system no longer applies straightforwardly, the results of Corollaries 4.1 and 4.2 still hold. We examine the communicative success rate, since this measure is equivalent to the impoverished signalling game previously discusses—0.10 at chance and 0.20 at a signalling system. After 10^6 plays per run, the communicative success rate is 0.1997, on average (1000 runs). Almost all runs (0.943) achieved a success rate greater than 0.1990. Learning is still fast.

When we introduce punishment into the impoverished signalling game, the sender and receiver fail to evolve a clear partition of nature more often than without punishment (this occurs in 0.250 runs). This is because the receiver now learns more quickly to *not* perform actions that often lead to failures. As was mentioned above, when this happens, the sender is equally punished for sending m_0 or m_1 in the state corresponding to the act which the receiver never chooses. Thus, over time, the signals in the urn for this state remain unbiased, and the sender signals randomly when this state is chosen.

Of the runs that evolve a clear partition, the unbiased $\langle 5, 5 \rangle$ partition is the most common of the 9 stable partitions for this game: again, about one third of the time 0.307, the sender and receiver perfectly partition nature for maximal information transfer. Most of the time (0.738), the sender and receiver partition nature *near*-perfectly—i.e., a $\langle 5, 5 \rangle$, $\langle 6, 4 \rangle$, or $\langle 4, 6 \rangle$ partition.

Figure 6 shows the distribution of runs for the impoverished signalling game with 10 state/act pairs and 2 signals under simple reinforcement as compared with the same parameters with punishment. Note again that the combinatorial measure does not perfectly track the actual distribution of partitions on simulation without punishment. However, when punishment is included in the learning dynamic, the distribution of partitions observed on simulation is somewhat more closely tracked by the combinatorial measure—though the distribution still tends to favour more informative partitions. This too is perhaps surprising, since the punishment is for miscoordination on *individual actions*—the sender and receiver are not punished for failing to evolve a maximally informative partition of nature.

Furthermore, learning is extremely sensitive to the parameters chosen. For example, when the sender receives both a reward and punishment $[+2, -0.1]$, but the receiver is not punished $[+1, 0]$, the distribution is qualitatively similar to those previously discussed, but the sender and receiver are more successful at evolving a clear partition of nature, failing only at a rate of approximately 0.04. When rewards are asymmetric in the other direction ($[+1, 0]$ for the sender, and $[+2, -0.1]$ for the receiver), they essentially always fail to evolve a clear partition (approximately 0.95). In these cases, the players still coordinate successfully for a success rate close to 0.20 after 10^6 plays per run. However, the receiver learns very quickly
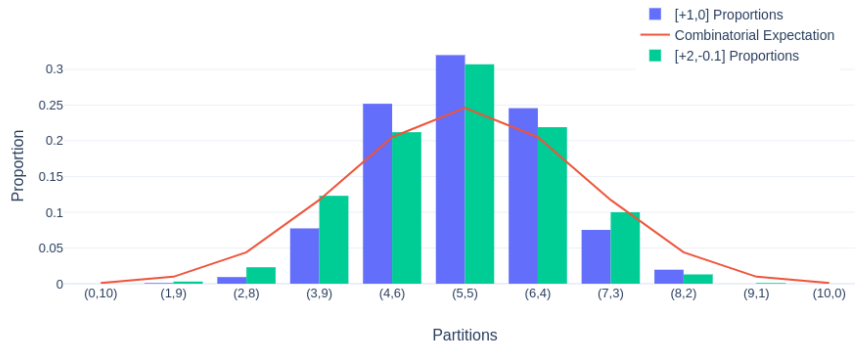
FIGURE 6. Partitioning 10 unbiased states with 2 signals. Comparison of impoverished signalling game with and without punishment

(with punishment) to simply ignore many of the actions—she puts all of her weight on a unique action for each of the signals. Thus, the sender never reinforces for the states corresponding to those actions; rather than learning to partition nature, she is learning to manipulate the (quickly fixed) dispositions of the receiver.[27]

This further serves to highlight the importance of the role asymmetries of the sender and receiver, as is discussed in Brusse and Bruner (2017); LaCroix (2019a). Note, that Brusse and Bruner (2017) suggest that signalling systems evolve more readily when the sender learns more quickly that the receiver, and that this result is quite robust—They highlight that Hofbauer and Huttegger (2008) show that signalling conventions are possible when the mutation rate of the a receiver population exceeds that of the sender population under the replicator-mutator dynamic. The results presented here are consistent with their analysis to the extent that they interpret this a situation in which the receiver (population) is relatively unresponsive to the sender population. This is precisely what happens in the impoverished signalling game when the receiver learns too quickly: the receiver becomes unresponsive to the sender, forcing the sender to learn how to 'manipulate' the receiver instead of learning a partition of nature.

5. Discussion

Barrett and LaCroix (2019) emphasise, as I have here, that a language which partitions nature into equally probable sets allows for maximal information transfer. We should certainly expect agents who are *rewarded* for communicating the most information per signal to evolve—for example, in terms of costly signalling. However, it has been demonstrated clearly that costly signalling is by no means a necessary condition for the evolution of such an efficient, maximally-informative language. Indeed, in many cases the sender and receiver evolve a nearly unbiased partition of nature under simple reinforcement learning, where signalling is cost-free and there is no punishment for miscoordination. Furthermore, when punishment is

---

[27]For a discussion of sensory manipulation in the context of signalling games, see Barrett and Skyrms (2017).

introduced for miscoordination, the players learn more often a maximally informative signalling system. Again, this is in spite of the fact that the cost of signalling is not a cost for being less informative.

Barrett and LaCroix (2019) use these results to explain how the structural properties of a language come to reflect the world in which the language evolved. This shows how something like a principle of indifference (in a Bayesian sense) might arise naturally in an evolutionary context. The key thing to note is that the naturalness of a particular partition depends inherently upon the context under consideration. With respect to communicative success, we have seen here that there is nothing that recommends a symmetric partition based upon the communicative success of a partition alone; nonetheless, individuals are more likely to choose partitions that maximise information transfer, given the communicative capacities with which they are endowed.

Payoff alone drives the dynamics under reinforcement learning: upon receiving a reward, the sender and receiver reinforce their behaviour, so reinforcement learning is driven by rewards. In spite of this analytic fact, we have seen that under simulation there is a natural tendency to move toward the most informative signalling system available. At first blush this should be somewhat surprising, given that every signalling system in the impoverished signalling game has equivalent communicative success. A priori, there is no reason to think that one type of partition will be favoured over another. In hindsight, however, we see that signalling systems are not in fact equiprobable, when accounting for distinct partitions, and the most likely signalling systems naturally come equipped with maximal information transfer. In fact, this point generalises to the extent that both the binomial coefficient and the quantity of information are functions that increase monotonically toward the unbiased partition (and decrease monotonically away from it).[28]

What I have shown in this paper is that agents learning to coordinate tend to favour maximal information transfer in spite of the fact that nothing from an initial analysis of the *stability properties* of the underlying signalling game suggests that this should be the case. Further, I have explained why this might be so—namely, the underlying structure of our model favours maximal information transfer in regard to the simple combinatorial properties of the ways in which the agents might partition nature. I showed that the structure of the game is such that there are more ways to achieve maximal information transfer than not for any particular partition. Furthermore, when nature is biased, the combinatorial argument alone cannot suffice to explain the observed frequency of the various partitions under simulation. This suggests that there is something over and above the combinatorial argument offered here that is causing the results that we see. However, the analytic connection between the partitions in the impoverished signalling game and maximal information transfer remains to be shown. This is an open question.

To the extent that the models presented accurately capture some real-world natural processes, we have seen that the evolutionary process is such that we can expect individuals to communicate as efficiently as possible (at least under the contexts which I have examined) and further that there is a natural tendency to achieve maximal information transfer. That being said, the models presented here are, of course, highly idealised and highly simplified—the actual world is significantly

---

[28]This follows in a straightforward way from the formula for entropy on the one hand and the formula for the binomial coefficient on the other.

more complex than the world which our sender and receiver inhabit. However, I suggest that they are illuminating nonetheless, in the very least in a 'how-possibly' sense. Further examination is clearly warranted. On a methodological note, then, this paper serves as an example to highlight the complementary roles of numerical simulations and analytic results.

References

Argiento, Raffaelle, Pemantle, Robin, Skyrms, Brian, and Volkov, Stanislav (2009). Learning to Signal: Analysis of a Micro-Level Reinforcement Model. *Stochastic Processes and Their Applications*, 119:373–390.

Barrett, Jeffrey (2006). Numerical Simulations of the Lewis Signaling Game: Learning Strategies, Pooling Equilibria, and Evolution of Grammar. *Technical Report, Institute for Mathematical Behavioral Science*.

Barrett, Jeffrey (2007). Dynamic Partitioning and the Conventionality of Kinds. *Philosophy of Science*, 74:527–546.

Barrett, Jeffrey (2009). The Evolution of Coding in Signaling Games. *Theory and Decision*, 67:223–237.

Barrett, Jeffrey (2016). On the Evolution of Truth. *Erkenntnis*, 81:1323–1332.

Barrett, Jeffrey (2017). Truth and Probability in Evolutionary Games. *Journal of Experimental and Theoretical Artificial Intelligence*, 29(1):219–225.

Barrett, Jeffrey and Zollman, Kevin (2009). The Role of Forgetting in the Evolution and Learning of Language. *Journal of Experimental and Theoretical Artificial Intelligence*, 21(4):293–309.

Barrett, Jeffrey A. and LaCroix, Travis (2019). The Co-Evolution of Language and Epistemic Principles. Unpublished Manuscript. May, 2019. PDF File.

Barrett, Jeffrey A. and Skyrms, Brian (2017). Self-Assembling Games. *The British Journal for the Philosophy of Science*, 68(2):329–353.

Barrett, Jeffrey A., Skyrms, Brian, and Cochran, Calvin (2018). Hierarchical Models for the Evolution of Compositional Language. Unpublished Manuscript. May, 2018. PDF File.

Barrett, Jeffrey A., Skyrms, Brian, and Mohseni, Aydin (2017). Self-Assembling Networks. *British Journal for the Philosophy of Science*. Forthcoming.

Birch, Jonathan (2014). Propositional Content in Signalling Systems. *Philosophical Studies*, 171(3):493–512.

Brusse, Carl and Bruner, Justin (2017). Responsiveness and Robustness in the David Lewis Signaling Game. *Philosophy of Science*, 84(5):1068–1079.

Donaldson, M. C., Lachmann, M., and Bergstrom, C. T. (2007). The Evolution of Functionally Referential Meaning in a Structured World. *Journal of Theoretical Biology*, 246:225–233.

Dretske, Fred (1981). *Knowledge and the Flow of Information*. The MIT Press.

Erev, Ido and Roth, Alvin E. (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *The American Economic Review*, 88(4):848–881.

Franke, Michael (2016). The Evolution of Compositionality in Signaling Games. *Journal of Logic, Language and Information*, 25(3):355–377.

Glimcher, Paul W. (2011). Understanding Dopamine and Reinforcement Learning: The Dopamine Reward Prediction Error Hypothesis. *Proceedings of the National Academy of Sciences*, 108(42):15647–15654.

Godfrey-Smith, Peter (2011). Signals: Evolution, Learning, and Information by Brian Skyrms (Review). *Mind*, 120(480):1288–1297.

Goodman, Nelson (1965). *Fact, Fiction, and Forecast.* The Bobs-Merrill Company, Inc., London.

Herrnstein, Richard J. (1970). On the Law of Effect. *Journal of Experimental Analysis of Behavior*, 13:243–266.

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

Hu, Yilei, Skyrms, Brian, and Tarrès, Pierre (2011). Reinforcement Learning in Signaling Game. *arXiv preprint arXiv:1103.5818.*

Huttegger, Simon M. (2007a). Evolution and the Explanation of Meaning. *Philosophy of Science*, 74:1–27.

Huttegger, Simon M. (2007b). Evolutionary Explanations of Indicatives and Imperatives. *Erkenntnis*, 66:409–436.

Huttegger, Simon M. (2007c). Robustness in Signaling Games. *Philosophy of Science*, 74(5):839–847.

Huttegger, Simon M., Skyrms, Brian, Smead, Rory, and Zollman, Kevin J. S. (2010). Evolutionary Dynamics of Lewis Signaling Games: Signaling Systems vs. Partial Pooling. *Synthese*, 172(1):177–191.

Huttegger, Simon M. and Zollman, Kevin J. S. (2011). Signaling Games. In Benz, A., Ebert, C., Jäger, G., and van Rooij, R., editors, *Language, Games, and Evolution*, volume 6207 of *Lecture Notes in Computer Science*, pages 160–176. Springer, Berlin, Heidelberg.

LaCroix, Travis (2019a). Accounting for Polysemy and Role Asymmetry in the Evolution of Compositional Signals. Unpublished Manuscript. May, 2019. PDF File.

LaCroix, Travis (2019b). Using logic to evolve more logic: Composing logical operators via self-assembly. Unpublished Manuscript. April, 2019. PDF File.

Lewis, David (2002/1969). *Convention: A Philosophical Study.* Blackwell, Oxford.

Roth, Alvin and Erev, Ido (1995). Learning in Extensive Form Games: Experimental Data and Simple Dynamical Models in the Intermediate Term. *Games and Economic Behavior*, 8:164–212.

Schultz, Wolfram (2004). Neural Coding of Basic Reward Terms of Animal Learning Theory, Game Theory, Micro-economics and Behavioural Ecology. *Current Opinion in Neurobiology*, 14(2):139–147.

Schultz, Wolfram, Dayan, Peter, and Montague, P. Read (1997). A Neural Substrate of Prediction and Reward. *Science*, 275:1593–1599.

Selten, Reinhard (1980). A note on evolutionarily stable strategies in asymmetric animal conflicts. *Journal of Theoretical Biology*, 84:93–101.

Shannon, Claude (1948). A Mathematical Theory of Communication. *The Bell System Mathematical Journal*, 27:379–423.

Shannon, Claude and Weaver, Warren (1949). *The Mathematical Theory of Communication.* University of Illinois Press, Urbana and Chicago.

Shea, Nicholas, Gofrey-Smith, Peter, and Cao, Rosa (2018). Content in Simple Signalling Systems. *The British Journal for the Philosophy of Science*, 69(4):1009–1035.

Skyrms, Brian (2010a). *Signals: Evolution, Learning, & Information.* Oxford University Press, Oxford.

Skyrms, Brian (2010b). The Flow of Information in Signaling Games. *Philosophical Studies*, 147(1):155–165.

Skyrms, Brian and Barrett, Jeffrey A. (2018). Propositional Content in Signals. *Studies in the History and Philosophy of Science C.* Forthcoming.

Steinert-Threlkeld, Shane (2016). Compositional signaling in a complex world. *Journal of Logic, Language, and Information*, 25(3–4):379–397.

Thorndike, Edward L. (1905). *The Elements of Psychology.* The Mason Press, Syracuse.

Thorndike, Edward L. (1911). *Animal Intelligence: Experimental Studies.* The Macmillan Company, New York.

Thorndike, Edward L. (1927). The Law of Effect. *American Journal of Psychology*, 39:212–222.

Trapa, Peter E. and Nowak, Martin A. (2000). Nash Equilibria for an Evolutionary Language Game. *Journal of Mathematical Biology*, 41(2):172–188.

Wärneryd, Karl (1993). Cheap Talk, Coordination and Evolutionary Stability. *Games and Economic Behavior*, 5(4):532–546.

Zollman, Kevin J. S. (2011). Separating Directives and Assertions Using Simple Signaling Games. *The Journal of Philosophy*, 108(3):158–169.

Appendix A. Extended Proofs

Proposition 3.1: The maximal communicative success rate for the impoverished signalling game, as defined in Definition 2.3, is

$$\max \pi(\sigma, \rho) = \frac{|M|}{|S|}.$$

Proof. The communicative success rate, as per Definition 2.1, is given by

$$\pi(\sigma, \rho) = \sum_{s \in S} P(s) \sum_{a \in A} u(s, a) \cdot \left(\sum_{m \in M} \sigma(s)(m) \cdot \rho(m)(a) \right)$$

Under the assumption that nature is unbiased, as per Definition 2.3, it follows that this is equivalent to

$$\pi(\sigma, \rho) = \frac{1}{|S|} \left[\sum_{a \in A} u(s, a) \cdot \left(\sum_{m \in M} \sigma(s)(m) \cdot \rho(m)(a) \right) \right]$$

Under the assumption that the utility function is as it is given in Definition 2.3, $u(s_i, a_j) = 0$ whenever $i \neq j$. Therefore, this further simplifies to

$$\pi(\sigma, \rho) = \frac{1}{|S|} \left[\sum_{i=1}^{|S|} \sum_{j=1}^{|M|} \sigma(s_i)(m_j) \cdot \rho(m_j)(a_i) \right]$$

Note that $\sigma(s_i)(m)$ is equivalent to $\mathbb{P}(m|s_i)$, and similarly $\rho(m)(a_i)$ is equivalent to $\mathbb{P}(a_i|m)$. Thus, we can expand the summation inside the brackets. For ease of exposition, I will represent this expanded sum as a table of values, with each cell containing an individual summand.

$$
\begin{bmatrix}
\mathbb{P}(m_1|s_1)\mathbb{P}(a_1|m_1) & \cdots & \mathbb{P}(m_{|M|}|s_1)\mathbb{P}(a_1|m_{|M|}) \\
\mathbb{P}(m_1|s_2)\mathbb{P}(a_2|m_1) & \cdots & \mathbb{P}(m_{|M|}|s_2)\mathbb{P}(a_2|m_{|M|}) \\
\vdots & \ddots & \vdots \\
\mathbb{P}(m_1|s_{|S|})\mathbb{P}(a_{|S|}|m_1) & \cdots & \mathbb{P}(m_{|M|}|s_{|S|})\mathbb{P}(a_{|S|}|m_{|M|})
\end{bmatrix}
$$

By Definition 3.1, a partition is a set of states such that $\mathbb{P}(m|s) = 1$. Therefore, by the law of total probability, it follows that exactly 1 the signal-probabilities conditional on the state equals 1 for each row of the summand written above, and the rest equal zero.

Thus, we write our sum more compactly as follows:

$$
\begin{bmatrix}
\mathbb{P}(a_0|m_0) & \mathbb{P}(a_0|m_1) & \cdots & \mathbb{P}(a_0|m_{|M|}) \\
\mathbb{P}(a_1|m_0) & \mathbb{P}(a_1|m_1) & \cdots & \mathbb{P}(a_1|m_{|M|}) \\
\vdots & \vdots & \ddots & \vdots \\
\mathbb{P}(a_{|S|}|m_1) & \mathbb{P}(a_{|S|}|m_2) & \cdots & \mathbb{P}(a_{|S|}|m_{|M|})
\end{bmatrix}
$$

I just noted that many of these cells will be equal to 0, since they will be multiplied by 0. However, what is important to note is that (1) This table has dimensions $|S| \times |M|$, and (2) all of the columns sum to 1, by the law of total probability. Therefore, our entire sum reduces to $|M|$ additions of 1.

Therefore, we have

$$
\pi(\sigma, \rho) = \frac{1}{|S|}\left[|M|\right]
$$

Therefore, under the assumption that all states are equiprobable, and the utility function is as in Definition 2.3, it follows that the maximal communicative success rate for this signalling game is

$$
\max \pi(\sigma, \rho) = \frac{|M|}{|S|}.
$$

$\square$

Corollary 3.1: In a signalling game, as defined in Definition 2.3, a strategy wherein the sender does not take advantage of all of the signals available to her has a lower communicative success rate than a strategy which does take advantage of all the signals available to her.

Proof. By Proposition 4.1, we know that the maximal payoff available to the sender and receiver is given by $\frac{|M|}{|S|}$. I note that if the sender does not use a particular signal, then the receiver will pool all her actions to the remaining $|M| - 1$ possible messages that she will receive. If she did not, then her strategy would not be a best response to the sender's strategy, σ. It follows, by the same argument as in Proposition 4.1 that the maximum expected payoff for a strategy of this sort will be

$$
\frac{|M| - 1}{|S|} < \frac{|M|}{|S|}.
$$

Therefore, the sender and receiver cannot acheive maximal payoff unless they take advantage of all the available messages. □