Inventing New Signals

Brian Skryms
University of California, Irvine
Stanford University

S. L. Zabell
Northwestern University

June 11, 2009

Abstract

A model of inventing new signals is introduced in the context sender-receiver games with reinforcement learning. If the invention parameter is set to zero, it reduces to basic Roth-Erev learning applied to acts rather than strategies, as in Argiento et. al. (2009). If every act is uniformly reinforced in every state it reduces to the Chinese Restaurant Process - also known as the Hoppe-Pólya urn - applied to each act. The dynamics can move players from one signaling game to another during the learning process. Invention helps agents avoid pooling and partial pooling equilibria.

Key words and phrases: Signals, Invention, Reinforcement, Chinese Restaurant Process, Hoppe-Pólya urn.

1 Introduction

Sender-receiver signaling games were first introduced by Lewis (1969) and then in more general form by Crawford and Sobel (1982). Nature picks a state of the world, with some fixed probability, from a set of states. One player, the sender, observes the state and picks a signal from some (arbitrary) set of signals. (Signals are not assumed to have any preexisting meaning or salience.) A receiver observes the signal and chooses one of a set of acts. Payoffs are jointly determined by state of the world and the act taken. It is interesting to investigate whether either the adaptive dynamics of evolution or learning can spontaneously generate meaningful signaling in such games.

Recent investigations reveal that unexpected complexity can be found in the dynamics of some very simple signaling games having strong common interest. Suppose, for example, that the sets of states, signals and acts are finite and of equal cardinality. Suppose further that for each state there is a unique act, such that the payoff for both sender and receiver is 1 if the act is performed in that state, and 0 otherwise. A sender's strategy is a map from states to signals, and a receiver's strategy is a map from signals to acts. If we give the unique act that pays off for a state the same index as that state, a *signaling system equilibrium* is a pair of sender and receiver strategies whose composition maps the i-th state to the i-th act: $s_i \rightarrow a_i$. There are also *complete pooling equilibria*, in which the sender sends the signals with some probability independent of the state and the receiver performs acts with a probability independent of the signal received, and (provided there are more than two states) *partial pooling equilibria* in which some but not all states are pooled.

In these special games, signaling system equilibria are distinguished from an evolutionary point of view by being the unique evolutionarily stable states (Wärneryd, 1993). It might therefore seem plausible that replicator dynamics would always lead to a signaling system equilibrium. And in the special case of 2 states, 2 signals and 2 acts with pure common interest (i. e., a common payoff), and where nature chooses states with equal probability, this is in fact the case. But it is not true in general (Hofbauer and Huttegger 2008, Pawlowitsch 2008). If the two states are not equiprobable, the connected component of pooling equilibria has a basin of attraction of positive measure. And with 3 states, 3 signals, and 3 acts, partial pooling equilibria have a basin of attraction of positive measure even if nature chooses states with equal probability.

With reinforcement learning, the situation is a little more complicated. Argiento et al. (2008) consider the reinforcement learning of Roth and Erev (1995) applied to 2 state, 2 signal, 2 act signaling games with pure common interest and equiprobable states. It generates a stochastic process as follows. The sender has an urn for each state. When she observes a state, she draws either a yellow or a blue ball from the urn for that state. If she draws a yellow ball she sends signal one; if she draws a blue ball she sends signal two. (Sender's urns initially have some yellow and some blue balls–it doesn't matter how many.) The receiver has an urn for each signal. Upon receiving a signal, the receiver draws a red or a green ball from the urn. If it is red he does act one; if it is green he does act two. (Receiver's urns initially have some red and some green balls.) Balls drawn are returned and if the right act for the state is done both sender and receiver add an extra ball of the color drawn to the urn drawn from. This is repeated. Using stochastic approximation theory, Argiento et. al. prove that with probability one the players converge to a signaling system equilibrium. If the initial distribution of colors is uniform, each signaling system equilibrium has equal probability of being selected (Theorem 1.1). But this result does not carry over to the case where states are not equiprobable. Sometimes there is convergence instead to a pooling equilibrium, just as with the replicator dynamics.

Why can't the agents simply *invent new signals* to remedy the situation? We would like to have a simple, easily-studied model of this process. That is to say, we want to move beyond closed models where the theorist fixes the signals, to an open model in which the space of signals can itself evolve. We would like to suggest one here. This involves a kind of hybrid of the Roth-Erev urn process and the Chinese restaurant process – the latter being also known in another guise as the Hoppe–Pólya urn.

2 Invention

The method used in this paper to introduce new signals has been studied in a variety of fields, including mathematical statistics, the theory of random permutations, and population biology. We begin by discussing two concrete, closely-related descriptions of it.

2.1 The Chinese Restaurant Process

Imagine a Chinese restaurant, with an infinite number of tables, each of which can hold an infinite number of customers. The basic mechanism is this: if n customers have already arrived, then the next customer sits to the left of each of the n already seated with probability $1/(n+1)$, and goes to a new table with probability $1/(n+1)$.

For example, the first person to arrive sits down at a table. The next person to arrive can now either sit down to the left of the first, or sit at an unoccupied table. Should the second join the first, the third to arrive has a 1/3 chance of sitting to the left of each of the previous two, and a 1/3 chance of sitting down at a new table. And so on. One can introduce a parameter $\theta > 0$ into the process by specifying that after n customers have arrived, the probability that the next customer sits to the left of each is $1/(n+\theta)$, and at a new table $\theta/(n+\theta)$.

This is the *Chinese restaurant process* (see, e. g., Aldous 1985, p. 92, Pitman 1996, Pitman 2006).[1] It provides a simple concrete representation of an important stochastic process that has been extensively studied for the last several decades, one that comes up in a surprising number of contexts.[2]

2.2 Hoppe's urn

If one disregards the seating arrangement of customers at a table, keeping track only of which table is occupied, then this simplified version of the Chinese restaurant process is equivalent to a simple urn model, one which is both a modification of Pólya's urn, and can be modified in turn to represent reinforcement learning with invention.

In a classical Pólya urn process, we start with an urn containing various colored balls. Then we proceed as follows: A ball is drawn at random. It is returned to the urn with another ball of the same color. Then repeat. All colors are treated in exactly the same way. We can recognize the Pólya urn process as a special case of reinforcement learning in which there is no distinction worth learning – all choices (colors) are reinforced equally. The probabilities in a Pólya urn process converge to a random limit. They are guaranteed to converge to something, but that something can be anything.

In 1984 Hoppe introduced what he called "Pólya-like urns" in connection with "neutral" evolution – evolution in the absence of selection pressure. To the Pólya urn, Hoppe (1984) adds a black ball – the *mutator*. The mutator

[1] The name has its genesis in the fact that in the 1970s, after the monthly Berkeley-Stanford Statistics colloquium, people would often meet for dinner at a local Chinese restaurant. The process is the brainchild of Jim Pitman and Lester Dubins, both members of the Berkeley Statistics Department. It first appeared in print in Aldous, 1985, although one of the authors of this paper remembers Jim Pitman describing it to him in 1979–1980.

[2] Here is one illustration. The seating arrangement after n customers arrive describes a permutation of the integers from 1 to n, with the customers at a table giving a cycle of the permutation. For example, suppose there are $n = 10$ customers sitting at three tables: customers $1 \to 9 \to 4 \to 1$ at table 1 (where $x \to y$ means that y is to the left of x), customers $2 \to 10 \to 6 \to 3 \to 2$ at table 2, and customers $5 \to 7 \to 8 \to 8$ at table 3. The corresponding permutation has the cycle decomposition $(1, 9, 4)(2, 10, 6, 3)(5, 7, 8)$. From this perspective the process randomly generates a permutation; and, in particular, if $\theta = 1$ it is not hard to see that the distribution is uniform. Random permutations have been intensively investigated, and the above representation is a useful tool in their study.

brings new colors into the game. If the black ball is drawn, it is returned to the urn and a ball of an entirely new color is added to the urn. (Hoppe allowed that there might be multiple black balls, corresponding to cases other than $\theta = 1$ in the Chinese restaurant process. Here, however, we will stick to the simplest case of one black ball.) The Hoppe–Pólya urn model was meant as a model for neutral selection, where there are a vast number of potential mutations which convey no selective advantage.[3]

It is evident that this urn process and the (simplified) Chinese restaurant process come to the same thing. Hoppe's colors correspond to the tables in the Chinese restaurant; the mutator ball corresponds to the possibility of sitting at a new table. After n iterations, the n guests in the restaurant or the n balls in Hoppe's urn are partitioned into some number of categories. The categories are tables for the restaurant, colors for the urn. But the partitions we end up with (who goes to which table, what color is selected) can be different each time; they depend on the luck of the draw.

2.2.1 Random partitions

Either way, each time the process is run we have a *random partition* of the integers 1, 2, ..., n, resulting in a different number of categories, different numbers of individuals in the categories, and different individuals filling out the numbers. The resulting patterns can be summarized by *partition vectors* $\langle a_1, ..., a_n \rangle$, where a_n is the number of tables with n guests. Suppose, for example, that six guests have been seated (so that one could have as many as six tables with one guest each, or as few a one table with six guests). Then the pattern of one table with three guests and three tables with one guest corresponds to the partition vector $a_1 = 3, a_3 = 1, a_2 = a_4 = a_5 = a_6 = 1$, or:

$$\langle 3, 0, 1, 0, 0, 0 \rangle \, .$$

It is easy to see that all $\binom{6}{3} = 20$ ways of realizing the pattern of one table with three guests and three tables with one have the same probability of occurring (1 in 192). It is generally true of the process that all realizations of a given partition vector are equally likely. All that affects the probability

[3]The same urn model has an alternative life in the Bayesian theory of induction, having essentially been invented in 1838 by the logician Augustus de Morgan to deal with the prediction of the emergence of novel categories. It generalizes Bayes-Laplace rules of succession, known to philosophers as Carnap's continuum of inductive methods (Zabell 1992, 2005).

is a specification of the number of tables and how many guests are seated at each. The property that any arrangement of guests with the same partition vector has the same probability is termed *exchangeability*, and is the key to mathematical analysis of the process.

There are explicit formulas to calculate probabilities and expectations of classes of outcomes after a finite number of trials. The expected number of categories – of colors of ball in Hoppe's urn or the expected number of tables in the Chinese restaurant – will be of particular interest to us, because the number of colors in a sender's urn will correspond to the number of signals in use. This expected number is given by a very simple formula:

$$E_n = \sum_{i=0}^{n-1} \frac{1}{1+i}.$$

Results are plotted in Figure 1.

Figure 1 goes here

It is known that *in the limit* as $n \to \infty$ the number of categories is almost surely infinite (that is, except for a set of sequences having probability zero).[4] Nevertheless even for large n the expected number of categories is relatively modest (because E_n, a partial sum of the harmonic series, grows logarithmically).

2.2.2 The Ewens sampling formula

There is something else that we would like to emphasize. For a given number of categories, the distribution of outcomes among those categories can be very nonuniform. (For a given number of tables, the number of customers from table to another can vary substantially.) Here is a simple example. Suppose there are ten trials and just two categories result (two colors of ball, two tables in the restaurant), which happens about 28% of the time. This can be realized in five different ways of partitioning 10:

$$5 + 5, \quad 4 + 6, \quad 3 + 7, \quad 2 + 8, \quad 1 + 9.$$

[4]This is an immediate consequence of the second Borel-Cantelli lemma, the displayed infinite series diverging. But exceptional sequences are *possible*, even if they form a set of probability zero: for example, every ball can have the same color as the first.

There is a simple formula for calculating the probabilities of these partitions – the *Ewens sampling formula* – which gives the probability of a partition vector for n draws:

$$Pr\,\langle a_1, ..., a_n \rangle = \prod_{j=1}^{n} \frac{1}{(j^{a_j})(a_j!)}$$

These probabilities in the case of our example are graphed in Figure 2.

Figure 2 goes here

Note that the more *unequal* a division is between the categories, the more *likely* it is to occur. Typically some colors are numerous, some are rare. Some tables are much fuller than others. This can be seen as the result of a kind of *preferential attachment* process. In the Chinese restaurant process, fuller tables are more likely to attract the new arrivals. (Who wants to be at a boring table?) This generates a so-call *power law distribution*, similar to those that are ubiquitous in word frequencies counts in natural language and elsewhere (Zipf 1932).

Finally, let us notice that the Hoppe urn can be redescribed in a suggestive way. Let us suppose that there are a lot of Pólya urns with different numbers of colors of balls, and different initial numbers of the balls. The mutator process is kept track of on the side, say with an urn with one black and many white balls. Pick a white ball from the auxiliary urn and you add another white ball, and sample from your current Pólya urn. Pick the black ball from the auxiliary urn and you move to a different Pólya urn with one more ball of one more color.

3 Reinforcement with Invention

We remarked that Pólya urn process can be thought of as reinforcement learning when there is no distinction worth learning – all choices (colors) are reinforced equally. The Hoppe-Pólya urn, then, is a model which adds useless invention to useless learning. That was its original motivation, where different alleles confer no selective advantage.

If we modify the Pólya urn by adding differential reinforcement – where choices are reinforced according to different payoffs – we get the basic Roth-Erev model of reinforcement learning, used by Argiento, et. al.. If we modify the Hoppe-Pólya model by adding differential reinforcement, we can get reinforcement learning that is capable of invention. See Figure 3.

Figure 3 goes here

3.1 Inventing New Signals

We use the Hoppe-Pólya urn as a basis for a model of inventing new signals in signaling games. For each state of the world, the sender has an additional choice: *send a new signal.* A new signal is always available. The sender can either send one of the existing signals or send a new one. Receivers always pay attention to a new signal. (A new signal means new signal that is noticed, failures being taken account of by making the probability of a successful new signal smaller.) Receivers, when confronted with a new signal, just act at random. We equip them with equal initial propensities for the acts.

Now we need to specify exactly how learning proceeds. Nature chooses a state and the sender either chooses a new signal, or one of the old signals. If there is no new signal the model works just as before with basic Roth-Erev reinforcement. If a new signal is introduced, it either leads to a successful action or not. When there is no success, the system returns to the state it was in before the experiment with a new signal was tried.

But if the new signal leads to a successful action, both sender and receiver are reinforced. The reinforcement now constitutes the sender's new initial propensity to send the signal in the state in which it was just sent. The receiver now begins keeping track of the success of acts taken upon receiving the new signal. In terms of the urn model, the receiver activates an urn for the signal, with one ball for each possible act, and adds to that urn the reinforcement for the successful act just taken. The sender now considers the new signal not only in the states in which it was tried out, but also considers it a possibility in other states. So, in terms of the urn model, a ball for the new signal is added to each senders urn, in addition to the reinforcement ball added to the urn for the state that has just occurred. The new signal has now established itself. We have moved from a Lewis signaling game with N signals to one with $N + 1$ signals. See Figure 4.

Figure 4 goes here

In summary, one of three things can happen, depending on whether a new signal is tried, and if it is successful:

1. A new signal is not tried; the game is unchanged. Reinforcement proceeds as in a game with a fixed number of signals.

2. A new signal is tried unsuccessfully; the game is unchanged.

3. A new signal is tried successfully, and the game changes from one with n states, m signals, a acts to one with n states, $m+1$ signals, a acts.

3.2 Starting with Nothing

If we can invent new signals, we can start with no signals at all, and see how the process evolves. We can expect that – like the simple Hoppe-Pólya urn – the limiting number of signals will be infinite. The appendix gives a proof that in the case of k states having unequal probability, k acts, and starting with no signals, the limiting number of signals is almost surely infinite.

We do not have at present anything like a stochastic approximation theory analysis of learning to signal with invention, and such an analysis looks very hard to come by. We therefore proceed to investigate the behavior of the process by simulation.

Consider the 3 state, 3 act Lewis signaling game with the states chosen with equal probabilities. As before, we have strong common interest – exactly one act is right for each state. In simulations of our model of invention, starting with no signals at all, the number of signals after 100,000 iterations ranged from 5 to 25. A histogram of the final number of signals in 1,000 trials is shown in Figure 5. This behavior is close to that which would be expected from a pure Chinese restaurant process.

Figure 5 goes here

3.3 Avoiding Pooling Traps

In a version of this signaling game with the number of signals fixed at 3, reinforcement learning sometimes leads to a partial pooling equilibrium. Simulations were carried out for basic Roth-Erev reinforcement learning, implemented as follows. The sender has 3 urns, one for each state. Each is initialized with one ball for each of the signals. The receiver has 3 urns, one for each signal. Each is initialized with one ball for each act. Nature chooses states with equal probability. On a successful round, both sender and receiver reinforce with one ball. In these simulations, about 9% of the trials led to imperfect information transmission. Using reinforcement with invention, starting with no signals, 1,000 trials *all* ended up with efficient signaling (defined as the probability of successful signaling in excess of 95% after 10^6 trials). Signalers went beyond inventing the three requisite signals. Lots of synonyms, signals that are sent in the same states and elicit the same acts, were created. By inventing more signals, one can avoid the traps of partial pooling equilibria.

In the game with 2 states, 2 acts, and the number of signals fixed at 2, if the states had unequal probabilities agents sometimes fell into a total pooling equilibrium – in which no information at all is transmitted. In such an equilibrium the receiver would simply do the act suited for the most probable state and ignore the signal and the sender would send signals with probabilities that were not sensitive to the state.

The probability of falling into a total pooling equilibrium increases as the disparity in probabilities increases. Simulations of basic Roth-Erev reinforcement learning illustrate this. In 1,000 trials, if one of the two states has probability $p = 0.6$, failure of efficient information transfer is rare (24/1000); this rises to 13% (132/1000) at $p = 0.7$, 26% (257/1000) at $p = 0.8$, and 49% (487/1000) at $p = 0.9$. Highly unequal state probabilities appear to be a major obstacle to evolution of efficient signaling.

If we take the extreme case in which one state has probability 0.9, start with no signals at all, and let the players invent signals as above, they reliably learn to signal. In 1000 trials they never fell into a pooling trap; they always learned a signaling system. They invented their way out of the trap. The invention of new signals makes efficient signaling a much more robust phenomenon.

3.4 Synonyms

Let us look at our results a little more closely. Typically we get efficient signaling with lots of synonyms. How much work are the synonyms doing? Consider the following trial of a 3 state, 3 act signaling game, starting with no signals and proceeding with 100,000 iterations of learning with invention (see Table 1).

Table 1 goes here

	Probability of Signal in		
Signal	State 0	State 1	State 2
1	0.00006	**0.71670**	0.00006
2	0.00006	**0.28192**	0.00006
3	0.09661	0.00006	0.00080
4	0.00946	0.00042	0.00012
5	**0.86867**	0.00012	0.00006
6	0.00006	0.00006	**0.81005**
7	0.02393	0.00006	0.00012
8	0.00006	0.00006	**0.14338**
9	0.00006	0.00018	0.04449
10	0.00012	0.00006	0.00043
11	0.00012	0.00012	0.00006
12	0.00054	0.00012	0.00018
13	0.00018	0.00006	0.00012

Table 1: Probability of signal in state

Notice that a few of the signals (shown in boldface) are doing most of the work. In state 1, signal 5 is sent 87% of the time. Signals 1 and 2 function as significant synonyms for state 2, being sent more that 99.5 % of the time. Signals 6 and 9 are the major synonyms for state 3. The pattern is fairly typical. Very often, many of the signals that have been invented end up little used.

This is just what we should expect from what we know about the Chinese restaurant process. Suppose n is the number of guests and L_r the numbers of guests at the r-th most populated table (so that $L_1 \geq L_2 \geq L_3 \geq \ldots$ and $L_1 + L_2 + L_3 + \ldots = n$). Shepp and Lloyd (1966, p. 348, Table 1) give the following values for $E[L_r/n]$, the expected fraction of guests at the r-th most populated table, in the limit as $n \to \infty$; see Table 2.

Table 2 goes here

Returning to our model, we see that even without any selective advantage, the distribution of reinforcements among categories tends to be very unequal, as shown in Figure 2. But there is still an element we have not yet captured in the model. Might not infrequently used signals simply fall out of use entirely?

r	$\lim_{n\to\infty} E[L_r/n]$
1	0.62432997
2	0.20958090
3	0.08831609
4	0.04034198
5	0.01914548
6	0.00927494
7	0.00454696
8	0.00224518
9	0.00111357
10	0.00055387
11	0.00027599
12	0.00013768
13	0.00006874
14	0.00003434
15	0.00001716

Table 2: Chinese restaurant process, limiting expected fraction of guests, r-th most populated table

4 Noisy Forgetting

Nature forgets things by having individuals die. Some strategies (phenotypes) simply go extinct. This can't really happen in replicator dynamics – an idealization where unsuccessful types get rarer and rarer but never actually vanish. And it can't happen in Roth-Erev reinforcement where unsuccessful acts are dealt with in much the same way.

Evolution in a finite population is different. In the models of Sebastian Shreiber (2001), a finite population of different phenotypes is modeled as an urn of balls of different colors. Successful reproduction of a phenotype corresponds to addition of balls of the same color. So far this is identical to the basic model of reinforcement learning. But individuals also die. We transpose the idea to learning dynamics to get a model of reinforcement learning with noisy forgetting.

For individual learning, this model may be more realistic than the usual model of geometrical discounting. That model, which discounts the past by keeping some fixed fraction of each type of ball at each update, may be best suited for aggregate learning where individual fluctuations are averaged out. But individual learning is noisy, and it is worth looking at an urn model of individual reinforcement with noisy forgetting.

4.1 Inventing and Forgetting Signals

We can put together these ideas to get learning with invention and noisy forgetting, and apply it to signaling. It is just like the model of inventing new signals except for the random dying-out of old reinforcement, implemented by random removal of balls from the sender's urns.

The idea may be implemented in various ways. Here is one. With some probability, nature picks an urn at random and removes a colored ball at random. (The probability is the forgetting rate, and we can vary it to see what happens.) Here is a second, suggested by Jason McKenzie Alexander. Nature picks an urn at random, picks a color represented in that urn at random, and removes a ball of that color.

Now it is possible that the number of balls of one color, or even balls of all colors could hit zero in a senders urn. Should we allow this to happen, as long as the color (the signal) is represented in other urns for other states? Here is another choice to be made. Preliminary simulations show that some forms of noisy forgetting are very effective in pruning unused signals, and others not at all. A more thorough investigation of these models will be pursued in future work.

5 Conclusion

We move from signaling games with a closed set of signals to more realistic models in which the number of signals can change. New signals can be invented, so the number of signals can grow. Little used signals may be forgotten, so the number of signals can shrink. A full dynamical analysis of these models is not yet available, but simulations suggest that such open models are more conducive to the evolution of efficient signaling than the previous closed models.

A Proof of the lower bound for growth of the number of signals

First some notation. Let $\mathcal{F}_n$ be the history of the process up to time n (the n-th trial). Let A_n denote the event that on the n-th trial a new signal is tried with success. Let ω denote the entire infinite history of a specific realization of our reinforcement process. (So if, instead, one were considering the process of tossing a coin an infinite number of times, ω would represent a specific infinite sequence of heads and tails.) Let

$$P(A_n \mid \mathcal{F}_{n-1})$$

denote the conditional probability that A_n occurs, given the past history of the process up to time $n-1$. This is not a number, but a random quantity, since it depends on the realization ω, which is random. Finally, let

$$P(A_n \mid \mathcal{F}_{n-1})(\omega)$$

denote this conditional probability for a specific history or realization ω; this is a number.

By the martingale generalization of the second Borel-Cantelli lemma (see, e. g., Durrett, 1996, p. 249), one has

$$\{A_n \text{ i. o.}\} = \left\{ \sum_{n=1}^{\infty} P(A_n \mid \mathcal{F}_{n-1}) = \infty \right\} \quad \text{almost surely.}$$

That is, consider the following two events. The first is

$$\{\omega : \omega \in A_n \text{ infinitely often}\};$$

the second event is

$$\left\{\omega : \sum_{n=1}^{\infty} P(A_n \mid \mathcal{F}_{n-1})(\omega) = \infty\right\}.$$

The assertion is that the two events are the same, up to a set of probability zero.

We claim that A_n occurs infinitely often with probability one; that is, $P(\{A_n \text{ i. o.}\}) = 1$. By the version of the Borel-Cantelli lemma just cited, it suffices to show that

$$P\left(\left\{\omega : \sum_{n=1}^{\infty} P(A_n \mid \mathcal{F}_{n-1})(\omega) = \infty\right\}\right) = 1.$$

In fact we show more: that

$$\sum_{n=1}^{\infty} P(A_n \mid \mathcal{F}_{n-1})(\omega) = \infty$$

for *every* history ω.

To see this, suppose that there are k states, and that the sender selects these with probabilities $p_j, 1 \leq j \leq k$. Suppose that initially there is just one ball, the black ball, in each of the k urns. Then at each stage there are $a_j \geq 1$ balls in each urn, one black and the remaining $a_j - 1$ some variety of colors. The probability that at stage n a new signal is generated and successfully used depends on the values of $a_1, ..., a_k$ at the start of stage n (that is, before selection takes place), and is

$$\sum_{j=1}^{k} p_j(\frac{1}{a_j})(\frac{1}{k}).$$

(That is, you pick the j-th urn with probability p_j, you pick the one ball out of the a_j that is black, and there is a one chance in k that the receiver chooses the correct act.)

Now use the generalized *harmonic mean–arithmetic mean inequality* (see, e. g., Hardy, Polya, and Littlewood, 1988, Chapter 2); this tells us that for $a_j > 0$, one has

$$\frac{1}{\sum_{j=1}^{k}\left(\frac{p_j}{a_j}\right)} \leq \sum_{j=1}^{k} p_j a_j.$$

Further, the total number of balls in the state urns, $a_1+...+a_k$, is greatest when a black ball has been selected every time and the receiver chooses the right act (since then one adds $k+1$ balls of a new color at each stage rather than just two). Thus at the start of stage n one has

$$a_1 + ... + a_k \leq (k+1)(n-1) + k$$

Thus if $p = \max\{p_j, 1 \leq j \leq k\}$, it is apparent that

$$\sum_{j=1}^{k} p_j a_j \leq p(a_1 + ... + a_k) \leq p\{(k+1)n - 1)\}.$$

Putting this all together gives us that the probability that at stage n a new signal is generated and successfully used is

$$\frac{1}{k}\sum_{j=1}^{k}\left(\frac{p_j}{a_j}\right) \geq \frac{1}{p}\frac{1}{(k+1)n-1}.$$

It follows that

$$\sum_{n=1}^{\infty} P(A_n \mid \mathcal{F}_{n-1})(\omega) \geq \frac{1}{p}\sum_{n=1}^{\infty}\frac{1}{(k+1)n-1} > \frac{1}{p(k+1)}\sum_{n=1}^{\infty}\frac{1}{n} = \infty,$$

using the well-known fact that the harmonic series diverges.

References

[1] Aldous, D. (1985) "Exchangeability and Related Topics", *École d'été de probabilités de Saint-Flour, XII –1983*, pp. 1–198. Berlin: Springer.

[2] 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* 119:373-390.

[3] Crawford, Vincent and Joel Sobel (1982) "Strategic Information Transmission" *Econometrica* 50:1431-1451.

[4] Durrett, Richard (1996) *Probability: Theory and Examples* 2nd ed. Belmont, CA: Duxbury Press.

[5] Hardy, G. H. Littlewood, J. E., and Pólya, G. (1988) *Inequalities* 2nd ed. Cambridge: Cambridge University Press.

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

[7] Hoppe, F. M. (1984) "Pólya-like Urns and the Ewens Sampling Formula" *Journal of Mathematical Biology* 20:91-94.

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

[9] Pawlowitsch, Christina (2008) "Why Evolution Does Not Always Lead to an Optimal Signaling System" *Games and Economic Behavior* 63:203-226.

[10] Pitman, J. (1995) "Exchangeable and Partially Exchangeable Random Partitions", *Probability Theory and Related Fields* 102: 145158.

[11] Roth, Al and Ido Erev (1995) "Learning in Extensive Form Games: Experimental Data and Simple Dynamical Models in the Intermediate Term" *Games and Economic Behavior* 8: 164-212.

[12] Shepp, L. A. and S. P. Lloyd (1966) "Ordered Cycle Lengths in a Random Permutation", *Transactions of the American Mathematical Society* 121:340–357.

[13] Shreiber, Sebastian (2001) "Urn Models, Replicator Processes and Random Genetic Drift" *SIAM Journal on Applied Mathematics* 61: 2148-2167.

[14] Wärneryd, Karl (1993) "Cheap Talk, Coordination, and Evolutionary Stability" *Games and Economic Behavior* 5: 532-546.

[15] Zabell, Sandy (1992) “Predicting the Unpredictable” *Synthese* 90:205-232.

[16] Zabell, Sandy (2005) *Symmetry and Its Discontents: Essays in the History of Inductive Probability* Cambridge: Cambridge University Press.

[17] Zipf, George Kingsley (1932) *Selected Studies in the Principle of Relative Frequency in the Language*. Cambridge, Mass.: Harvard University Press.