Sifting the Signal from the Noise

Daniel A. Herrmann & Jacob VanDrunen

Abstract

Signalling games are useful for understanding how language emerges. In the standard models the dynamics in some sense already knows what the signals are, even if they do not yet have meaning. In this paper we relax this assumption, and develop a simple model we call an 'attention game' in which agents have to learn which feature in their environment is the signal. We demonstrate that simple reinforcement learning agents can still learn to coordinate in contexts in which (i) the agents do not already know what the signal is and (ii) the other features in the agents' environment are uncorrelated with the signal. Furthermore, we show that, in cases in which other features are correlated with the signal, there is a surprising trade-off between learning what the signal is, and success in action. We show that the mutual information between a signal and a feature plays a key role in governing the accuracy and attention of the agent.

*This paper is forthcoming in the British Journal for the Philosophy of Science. Article DOI: `https://doi.org/10.1086/720805`

1 Introduction

Lewis-Skyrms signalling games are useful for understanding how simple communication systems emerge (Lewis [1969]; Skyrms [2010]). The idea is to set up a situation in which either (i) individuals learn to communicate over time through some kind of learning procedure, or (ii) a population with different strategies for responding to signals changes over time based on the success of the strategies. In (i), the agents need a specific protocol for sending and receiving signals, such that they can learn from the successes and failures of their endeavours. For (ii), the strategies encode specific dispositions an agent has for sending signals conditional on observations, and for conditioning their acts on the signals they receive.

In both cases, the traditional account makes assumptions about the agents' knowledge of the structure of the game. First, it assumes that agents know the space of possible actions and signals. But it also assumes that the agents already know in some sense what part of the world constitutes the signal. A thoroughgoing empiricism must be able to give an account of how these features of the game themselves arise from more fundamental adaptive processes. Alexander et al. ([2012]) give a model in which agents invent new signals over time. This provides an account of how agents might learn the space of possible actions. The collection of models we provide here addresses the other problem: how might agents learn which available stimulus is best to condition their actions on. This is a fundamental concern for the theory of self-assembling games.[1]

In a signalling game, when a sender sends a signal to her partner, the partner responds to that signal in a specific way. But in the real world, her partner might not know what part of the act the sender performs is meant to be the signal. Suppose she signals by waving a red flag. What is the signal here? Is it the colour? The fact that it is a flag? The pattern in which she waves it? Where she stands when she is waving it?

Our work draws inspiration from other models in the literature. Herrmann and Skyrms ([forthcoming]) provide a model of the invention of conventions in which agents need to learn the properties on which they condition their strategies. In the epistemic network game of Barrett et al. ([2019]), agents learn to attend to other agents and use an attention urn for this purpose. Relatedly, Bala and Goyal ([2000]) consider a model of network formation in which agents choose the other agents with whom they form costly connections, based on possible benefits. Barrett ([2020]) gives a Lewis-Skyrms signalling game model in which agents must learn to distinguish between a signal with an already-established meaning and an uncorrelated feature of the world when deciding how to condition their actions. He calls this process of appropriating already-evolved signals as inputs to a game *modular composition*.[2] Our concern here is with a more general form of modular composition. Instead of supposing that the sender has fixed signalling dispositions, we consider the case in which both

[1] See (Barrett and Skyrms [2017]) for the details of this theory.

[2] Lacroix ([2020b]) gives a related model of modular composition in which agents learn to make use of the pre-evolved dispositions of other agents instead of evolving them from scratch.

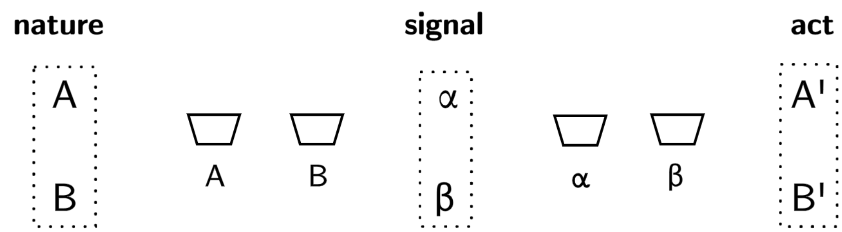

Figure 1: A $2 \times 2 \times 2$ Lewis-Skyrms signalling game.

senders and receivers must co-evolve a signalling system (as in the traditional Lewis-Skyrms signalling game). We then show how in the context of such a signalling game, the receiver might additionally learn to pick out the sender's signal from a set of possible stimuli.[3]

2 Lewis-Skyrms Signalling Games

Lewis ([1969]) proposed the signalling game to show how agents might use pre-existing saliences to coordinate on a conventional communication system. Nature presents some stimulus which yields a payoff if a particular action is taken in response. The sender observes the stimulus, and sends some signal to the receiver. The receiver in turn observes the signal and performs some action. If the action corresponds to the state of nature, the agents receive the payoff. Coordination in a traditional Lewis signalling game happens by means of prior agreement between players, salience, or precedent. Skyrms ([2010]) showed how coordination could evolve without appeal to any mechanism beyond simple adaptive processes.

As a simple example of this, consider the 2-state, 2-signal, 2-action ($2\times2\times2$) signalling game under a simple reinforcement learning dynamics.[4] An illustration of this game is shown in Figure 1. Nature presents one of two possible stimuli, uniformly at random. The sender draws from one of two urns corresponding to the stimulus. The ball has one of two possible colors, corresponding to the possible signals that he could send. He then sends the corresponding signal. The receiver observes the signal and draws from one of two urns corresponding to the signal. These urns in turn have balls of two colors, corresponding to the two possible actions. She then performs the corresponding action. If the action corresponds to the state of nature, the receiver returns the ball she drew to the urn, and adds another ball of that color. The sender does the same. If the

[3]In a signalling game, the signal might be any feature of the world which the sender has control over. What we are doing here is relaxing the requirement that the receiver know which feature of the world that is. There may be many features of the world salient to the receiver which the sender has only partial control over, or no control over whatsoever. The receiver must learn to distinguish the signal from the noise.

[4]This is a special case of a learning process commonly used in psychology and economics (Luce [1959]; Herrnstein [1970]; Erev and Roth [1998]), but it is also an evolutionary process in that agents are relying on an adaptive dynamics to carry them to equilibrium.

action does not correspond to nature, they return their balls to their respective urns, but do not add an additional ball.

For the case of $2\times2\times2$ signalling games under simple reinforcement, Argiento et al. ([2009]) proved that convergence to an optimal signalling system is guaranteed in the limit. This means that, in this simple context, agents will always learn a signalling system.

3 Attention Games

Here we introduce the attention game in abstract. Suppose we have a sender and a receiver. The sender makes some observation from an observation partition $N = \{\sigma_0, \sigma_1, \ldots, \sigma_n\}$. The sender then chooses which signal to send from a collection of possible signals $S = \{s_0, s_1, \ldots, s_n\}$.

Here is where things differ from the traditional Lewis signalling game. Instead of simply observing the signal and then acting, the receiver instead observes a *feature vector*, which is probabilistically generated from the signal. More formally, we define a *feature* f as a finite set of natural numbers, and a feature set $\mathcal{F} = \{f_0, f_1, \ldots, f_l\}$ as a set of features. A feature vector is an $l+1$-length vector such that its ith member v_i is a member of f_i.

The intended interpretation of these features is the following. We can think of the values that a feature can take as partitioning the possible observations of the receiver. For example, in the traditional $2 \times 2 \times 2$ Lewis signalling game, the receiver observes whether one signal or the other is sent. Since it is assumed one and only one signal is sent, the two signals together partition the receiver's possible observations. However, the receiver may pay attention to other partitions of their possible observations. These partitions may or may not relate to the signal partition.

Our features capture this idea of an observation partition. The set of possible feature vectors represents the possible observations that the receiver might make. The range of each feature partitions this set.[5]

Instead of the receiver simply conditioning their act $a \in \mathcal{A} = \{a_0, a_1, \ldots, a_n\}$ on the value of f_0, the receiver first must choose which feature to condition on when selecting their act. Nature then determines the payoffs according to whether or not the act matched the observation. An attention game differs from a signalling game by introducing on the receiver's side the choice to pay attention to different features, introducing the problem of distinguishing the signal from the noise.

4 Uncorrelated Features

To make this more concrete, consider a $2 \times 2 \times 2$ attention game under simple reinforcement learning. Suppose that there are 2 possible observations which

[5] We will later put probabilities over the values that features might take. Thus, we require every feature to be a measurable function, with respect to whatever background field. Then features are random variables. Since everything is finite here, it is clear that the measureability constraint poses no problem.

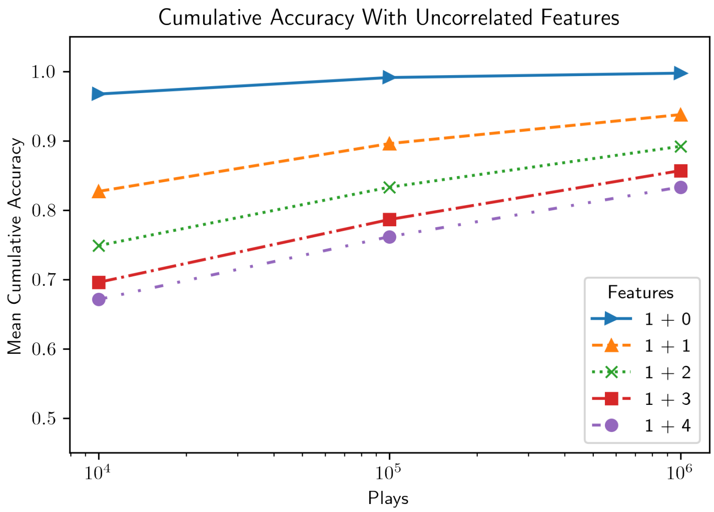

Figure 2: Mean cumulative accuracy for different run lengths (each data point represents the mean of 10^3 simulations). To emphasize the number of uncorrelated features i, we write $1 + i$, with 1 representing the signal feature.

nature will randomly determine with equal chances on each round, 2 possible acts corresponding to the states of nature, and 2 possible signals. Suppose that the number of features is at least 1, and for each i, $f_i = \{0, 1\}$. f_0 is the signal-feature, and all other f_i are determined randomly. All features are hence uncorrelated with each other.

Upon making an observation, the sender draws from the corresponding urn. The sender then sends the signal corresponding to the drawn ball. f_0 then takes on a value corresponding to the signal, and the other features take on values determined uniformly at random. The receiver then draws a ball from their attention urn, and observes the feature corresponding to the drawn ball. Finally, the receiver draws a ball from the urn corresponding to the chosen feature and the chosen feature's value, and takes the corresponding act. Note that the receiver has a unique urn for each pair of feature and feature value. If the action corresponds to the observation, the players are rewarded, and they return all drawn balls to their respective urns, and add an additional ball of the same type. If the action was unsuccessful, the players simply return all drawn balls to their respective urns.

The addition of the various features and the attention urn means that we cannot apply the Argiento et al. ([2009]) convergence results to the attention game. Here we will be giving simulation results which estimate the medium-run performance of learning agents in our model.

Results for this game under varying parameters are shown in Figure 2. In all versions, the game is set up so that one feature reflects the signal, and we vary the number of additional uncorrelated features from 0 (which would be a traditional $2 \times 2 \times 2$ Lewis-Skyrms signalling game) up to 4, with 10^3 simulations of each experimental condition. We see that the cumulative accuracy decreases significantly as the number of non-signal features increases. To use another metric, after 10^6 plays with 0 non-signal features, 0.996 of the runs ended up with cumulative accuracy above a threshold of 0.75.[6] With 4 non-signal features, this happens on only 0.708 of the runs. Learning is still possible, but the addition of "noise" in the form of uncorrelated features slows down the process.[7]

5 Correlated Features

We saw above that learning is frequent, even if slower, when non-signal features are uncorrelated with the signal-feature. But what happens when those features are imperfectly correlated with the signal-feature?

We consider here a particular kind of correlation between non-signal features and the signal-feature. Specifically, we consider cases in which a non-signal feature value depends on the signal to some extent, but not directly on the state of nature.[8] For simplicity we restrict our analysis to the case where features are all binary. We represent the extent to which a non-signal feature f is correlated with the signal-feature with two parameters, α and β, where

$$\alpha = P(f = 1 \mid f_0 = 1)$$
$$\beta = P(f = 1 \mid f_0 = 0)$$

These two parameters also characterize the unconditional probabilities that f takes on a particular value, as long as the probabilities of the signals are known. Note that when $\alpha = \beta$ the non-signal feature and the signal-feature are uncorrelated.[9]

6 Learning Results for Correlated Features

In order to understand the relationship between correlation and learning we consider an attention game with only one additional feature.[10] For ease of exposition, we refer to this non-signal feature as 'the feature', and denote it

[6] This threshold is chosen arbitrarily. Really, any threshold significantly higher than 0.5 will be good enough to show learning.

[7] Learning a signalling system becomes significantly harder when we extend the model such that the sender must also pay attention to the correct source among several uncorrelated sources in nature. Specifically, after 10^6 rounds, the agents in this setup only do better than 0.75 on 0.680 of the runs when 1 uncorrelated source of nature is added to the $1 + 1$ game. This drops to 0.141 when 4 uncorrelated sources of nature are added to the $1 + 4$ game. We thank Jeff Barrett for suggesting this extension to the model.

[8] Of course, if the feature is correlated with the signal-feature, and the signal-feature with the state of nature, then the feature is correlated with the state of nature. What we mean here is that the signal screens off this correlation. This could be relaxed in future work.

[9] We will soon use α and β to make precise the extent of correlation between feature and signal-feature. For this we use mutual information.

[10] So two features total: the signal-feature and one non-signal feature.

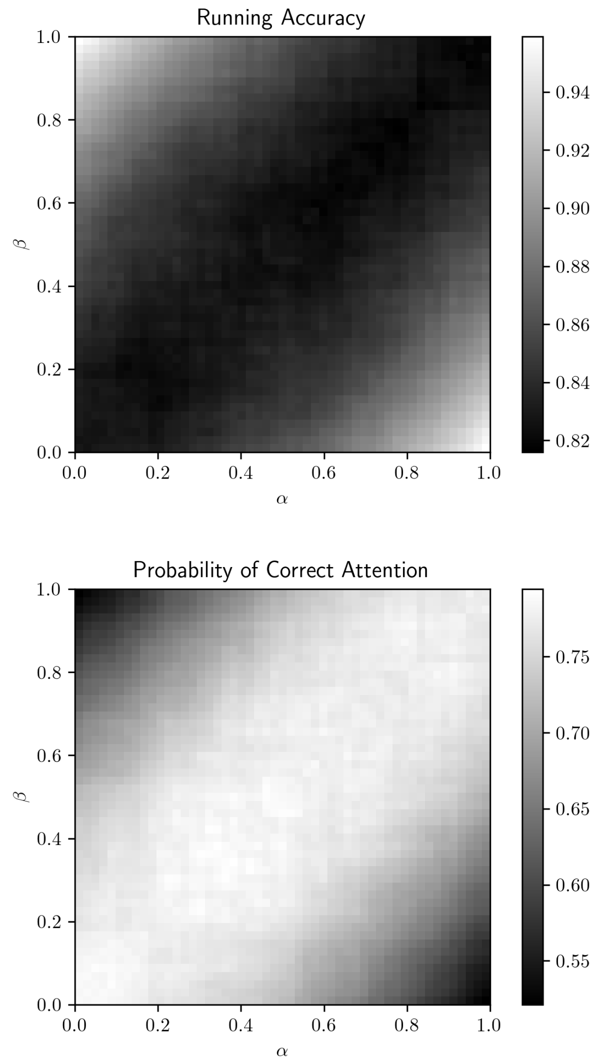

Figure 3: Running accuracy (top) and probability of correct attention (bottom) for experiments with one correlated feature. Heatmaps are mirrored from upper left to bottom right.

with f. We will use 'A' to denote the event $f = 1$, and s_1 to denote the event $f_0 = 1$.

The experimental conditions now depend on α and β. We sample α at intervals of 0.02 on $[0, 1]$, and, for every value of α, we sample β at intervals of

0.02 on $[0, \alpha]$.[11] We run 10^3 simulations of each condition, with 10^4 plays per simulation.

Results for this experiment are given in Figure 3. Here we observe that as the difference between α and β decreases – that is, as the correlation of the feature with the signal-feature decreases – the accuracy of the learned signalling convention decreases correspondingly. This means that agents spend less time paying attention to the most informative feature: the signal feature. On the other hand, the probability that the receiver draws the correct feature from their attention urn increases under the same conditions. This reveals a surprising trade-off. We might think that groups which tend more often to pay attention to the correct feature (the signal-feature) would be more successful at learning to perform the correct action. This is the opposite of what we observe.[12] In order to understand this trade-off better, we turn to a discussion of mutual information.

7 Mutual Information Between Signal and Feature

We use mutual information in order to characterize the extent to which a non-signal feature and a signal-feature are correlated.[13] The mutual information of two discrete random variables X and Y is given by

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}.$$

Intuitively, this quantity represents the amount of information one gets about one random variable if one observes the value of the other. This is a natural quantity for our context, in which we want to characterize how informative a feature is about a signal-feature.

The mutual information depends on the unconditional probabilities of the two random variables. In our context these change as the agents learn up a signalling convention. Thus, the mutual information will vary over time. To see this, consider the first term of the sum:

$$P(A, s_1) \log \frac{P(A, s_1)}{P(A)P(s_1)}.$$

Notice that this depends on the value $P(A, s_1)$, which we can rewrite as

$$P(A \mid s_1)(P(s_1 \mid \sigma_1)P(\sigma_1) + P(s_1 \mid \sigma_0)P(\sigma_0)) = \frac{\alpha}{2}(P(s_1 \mid \sigma_1) + P(s_1 \mid \sigma_0)).^{14}$$

[11] No practical difference is made by switching the values of α and β, so no further conditions are needed.

[12] One might worry about the cumulative accuracy not accurately reflecting the final outcome of learning. We also ran the same simulations with 10^3 non-learning plays at the end to determine a measure of the final accuracy at the conclusion of the learning process. The results are comparable in all relevant ways.

[13] See Skyrms ([2010]) for a discussion of the application of information theory to signalling games.

[14] The $\frac{1}{2}$ in the expression on the right comes from the equiprobable states of nature.

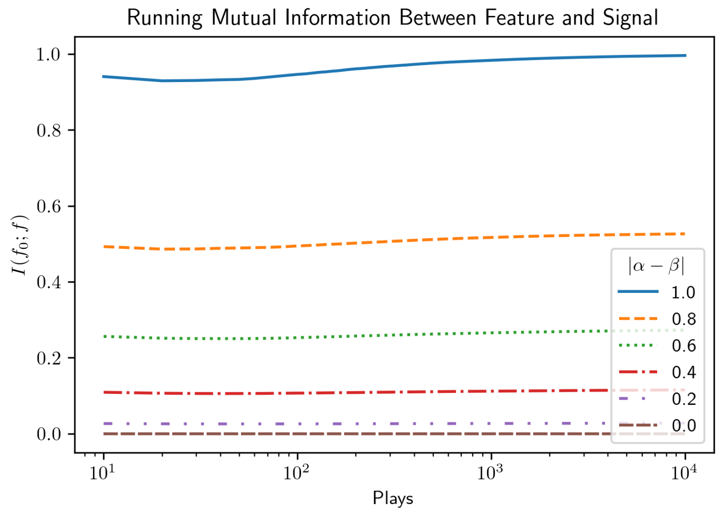

Figure 4: Running mutual information between feature and signal for various degrees of correlation.

The two conditional probabilities on the right hand side, $P(s_1 \mid \sigma_1)$ and $P(s_1 \mid \sigma_0)$, are what change as the agents learn a signalling system. However, notice the following. At the beginning of an attention game, both of these values are $1/2$, and thus sum to 1. Furthermore, once the agents have a signalling system in place, they will also sum to 1. Thus, even though these quantities do fluctuate over time, we should expect them only to deviate a bit in the beginning, and then get very close to 1 again as the agents keep learning. Since the mutual information only changes as these change, we should expect the mutual information to be relatively constant as well. Indeed, as can be seen in Figure 4, simulation results of the mutual information over time bear this out.

Whenever the sum of $P(s_1 \mid \sigma_1)$ and $P(s_1 \mid \sigma_0)$ equals 1, we can calculate the mutual information between the signal-feature and the feature as

$$I(f_0; f) = \frac{\left(\alpha \log \frac{2\alpha}{\alpha+\beta} + \beta \log \frac{2\beta}{\alpha+\beta} + (1-\alpha) \log \frac{2(1-\alpha)}{2-\alpha-\beta} + (1-\beta) \log \frac{2(1-\beta)}{2-\alpha-\beta}\right)}{2}.$$

In Figure 5 we calculate this value for pairs of parameters, $\alpha, \beta \in [0, 1]$.

Notice the striking similarity this heat map has to the two heat maps in Figure 3. When we test for the correlation between mutual information and both accuracy and attention, we get almost perfect correlation, specifically $r = 0.991$ between accuracy and mutual information, and -0.988 between attention and mutual information.

This strong correlation points to an explanation of the results we stated above. As the mutual information increases, mistakes in which the receiver

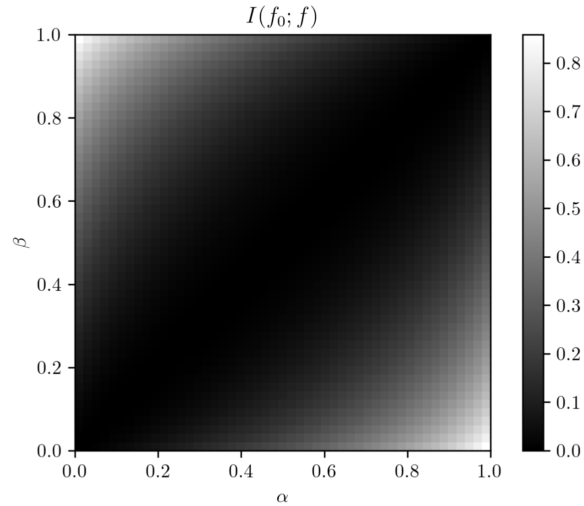

Figure 5: Heatmap of mutual information, measured with intervals of 0.02 for α and β.

pays attention to the less informative feature become less costly.[15] Thus overall accuracy increases. Conversely, as the mutual information increases, this also means that the non-signal feature is reinforced more often in the attention urn, which means that the receiver pays more attention to the less informative feature.[16] It is the mutual information that is driving these results. Once we know the mutual information of the feature and the signal-feature, we do not need to know the particular α and β parameters to predict what will happen.

[15]This is because the receiver is learning how to extract the information about the signal, and thus the state of nature, from the correlated feature.

[16]Here we have only given simulation results that show this to be the case in the medium-term. Whether the receiver persists in paying attention to non-signal features in the limit is an open question. Beggs ([2005]) proves analytically that in a game with a dominant strategy, simple reinforcement learning will learn to play that strategy with probability 1 in the limit. Showing which strategy dominates in the attention-learning process is subtle. Proving that paying attention to the signal-feature always dominates in the limit is equivalent to proving that the agents always learn a signalling system (separating equilibrium) in the limit. As previously mentioned, because the proof in (Argiento et al. [2009]) does not extend to cases in which the signalling channel is being manipulated by an outside process (in this case, the attention process), we cannot at this point extend the existing results to show that this will always be the case. When the agents are doing better than chance in the signalling game, however, there is some evidence that they will end up in the separating equilibrium in the limit.

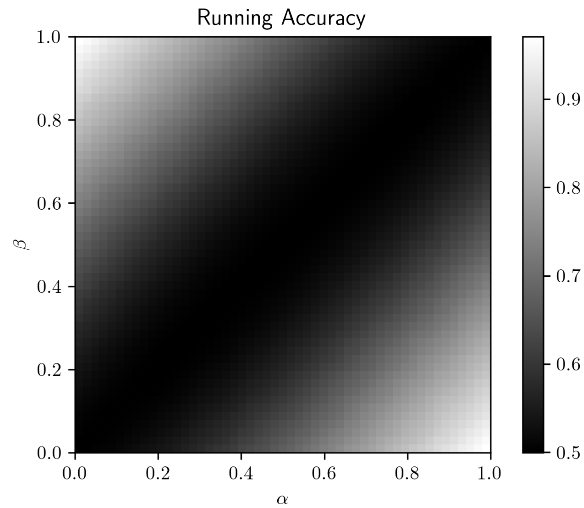

Figure 6: Heatmap of running accuracy (average across 10^3 simulations run for 10^4 plays) for experiments with no signal-feature and one correlated feature.

8 Features With No Signal-Feature

A natural thing to investigate is what happens if the signal-feature is not one of the features, and the features correlate to various degrees with the signal-feature. This attention game would be appropriate for modelling a situation in which the sender has only imperfect control over what the receiver gets to observe. In the case with only one feature this corresponds to a noisy signal-feature. In the case with multiple features, the receiver would have to learn which feature is most informative, and pay attention to it.

First we consider the case in which there is only one feature, which is not the signal-feature. In this case there is no attention urn.[17] We show the results in Figure 6.[18] Once again, mutual information is very highly correlated with accuracy ($r = 0.984$). This is what we should expect: the more informative a feature is about the signal-feature, the more successful the agents will be.

Finally, we consider the case where there are multiple features of differing correlation with the signal-feature, but no signal-feature. We want to understand both how adding more features affects accuracy, as well as the extent to which the receiver is able to pay attention to the more informative features. In

[17]This simple case might also be characterized as a signalling game with a noisy channel. Barrett et al. ([2017]) consider a similar case in the context of different learning dynamics and a different way of adding noise. Our noise is characterized by the α and β parameters, whereas in their model there is a small fixed probability of a random signal.

[18]Figure 6 looks smooth compared to those in Figure 3. We confirmed that the Figure 6 results are indeed showing the accuracy – the noise simply averaged out in this case because the attention-learning process has been removed.

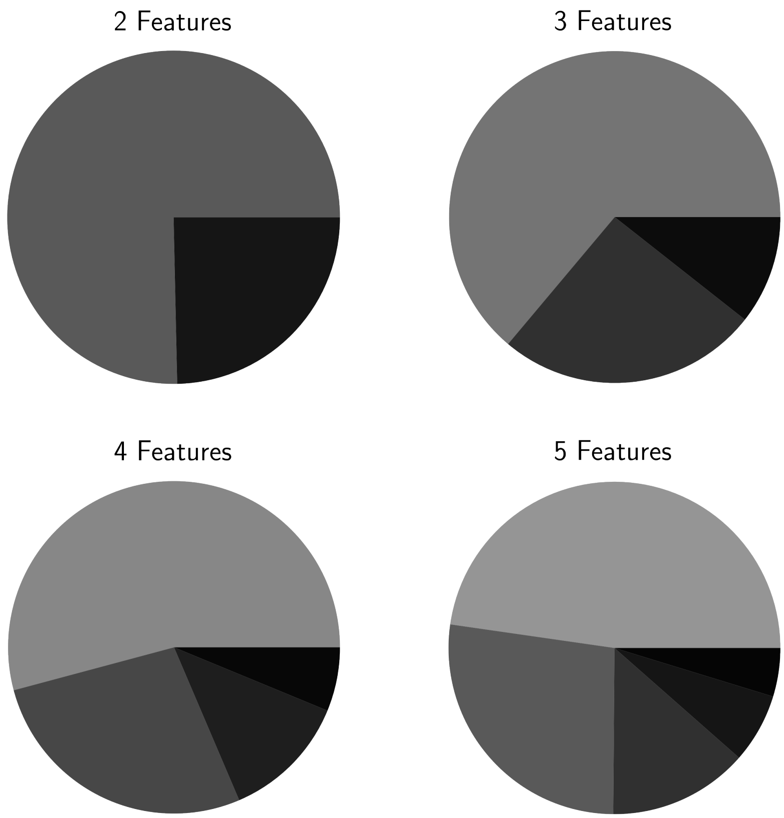

Figure 7: Pie charts showing the proportion of attention given to different features when no signal-feature is available. Lighter shading indicates a higher correlation of the feature (in terms of mutual information) with the unobservable signal-feature.

order to do this we consider four cases, where the number of features ranges from two to five. To generate the α and β parameters for each feature, we use the line segment $\overline{(0.5, 0.5), (1, 0)}$. For the case with n features, we divide this line segment into $n + 1$ sections, and take the coordinates of the dividing points as the values for α and β. This allows us to cut across the different possible values for the mutual information of the signal-feature and the feature in a way that maximizes the difference of the mutual information between each feature.

In the 2-feature condition, the mean cumulative accuracy was 0.766. For 3 features, 0.790. For 4, 0.797, and for 5, 0.806. The increasing accuracy observed in this experiment is due to the availability of more informative features as our partition of the line segment becomes finer. So, in the 2-feature case, the most informative feature is $\alpha = 0.83, \beta = 0.17$, while in the 5-feature case, it is $\alpha = 0.92, \beta = 0.08$. The proportion of balls corresponding to different features is shown in Figure 7. In each pie chart, lighter shading indicates a higher correlation of the feature (in terms of mutual information) with the signal-feature. We can see that the proportion of balls tracks the informativeness of the features, which accounts for the increase in accuracy.

9 Conclusion

In order for signalling games to explain the emergence of communication systems, they need to apply in less idealized situations. We considered a situation in which the receiver has to learn which one of many observable features is the signal-feature (if there is one). Indeed, insofar as the 'receiver' in a signalling game is one who receives a signal, what is of interest in this situation is how the role of being a receiver might itself emerge. We provided a model, the attention game, that captures this situation.

We showed that learning can still take place when there are multiple features that are uncorrelated with the signal-feature, even if it is slower. In the context in which there is one signal and one feature that may or may not be correlated with the signal-feature, we discovered a surprising trade-off between accuracy and attending to the signal-feature. We showed that the mutual information of the feature and the signal-feature is highly predictive of both accuracy and attention, and used this to propose an explanation for the accuracy-attention trade-off.

Finally, we considered cases in which the signal-feature was not one of the observable features. This is an important case since, in the real world, agents will never observe the signal of other agents with perfect fidelity. We showed that in the case with only one feature, mutual information once again predicted the success of the agents. In the case with multiple features of varying amounts of information, the receiver learns to pay attention to more informative features.

One potentially promising line of inquiry is to carry out a similar analysis and vary the number of states, acts, and signals.[19] Here we have demonstrated how the simplest $2 \times 2 \times 2$ Lewis-Skyrms signalling game might self-assemble by an attention-learning process. It would certainly be fruitful to investigate how the added attention-learning mechanism provides one path toward the self-assembly of these more complex signalling games as well. This would provide a precisification of the broad argument given in Barrett and Skyrms's original paper on self-assembling games ([2017]).

Another attractive line of inquiry would be the introduction of initially biased attention urns. This would be one way to incorporate Lewis's original category of salience into the Skyrmsian learning model, and to understand better the extent to which salience can decrease or increase the quality of learning.[20]

[19] One could go further and consider cases in which the different features have different numbers of possible values. This would of course make it more challenging to characterize the mutual information of the various features in a transparent way.

[20] See (LaCroix [2020a]) for a recent study of the effect a 'salience parameter' has on learning in Lewis-Skyrms signalling games.

Finally, our analysis focused on the learning dynamics of individuals. One might consider the attention games introduced here in an evolutionary context, and carry out the appropriate analyses. We doubt that the broad lessons we drew here would change dramatically, given the deep formal connections between the urn learning we consider here and the replicator dynamics often used in evolutionary analysis.[21] However, it would still be worthwhile to confirm this intuition and see if any other lessons emerge, or, even better, discover that it is false.

Acknowledgements

We would like to thank Jeffrey Barrett, Brian Skyrms, Benjamin Genta, and Gerard Rothfus for reading earlier drafts of the paper. We thank two anonymous referees for their thoughtful suggestions. The members of the LPS Game Theory Lab Group discussed an early draft of the paper. Finally, we thank the attendees of the 2022 ICLR Emergent Communication Workshop for thoughts on a condensed manuscript.

Daniel A. Herrmann
Department of Logic and Philosophy of Science
University of California, Irvine
Irvine, CA, USA
daherrma@uci.edu

Jacob VanDrunen
Department of Logic and Philosophy of Science
University of California, Irvine
Irvine, CA, USA
jvandrun@uci.edu

References

Alexander, J. M., B. Skyrms, and S. L. Zabell (2012). Inventing new signals. *Dynamic Games and Applications 2*(1), 129–145.

Argiento, R., R. Pemantle, B. Skyrms, and S. Volkov (2009, February). Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications 119*(2), 373–390.

Bala, V. and S. Goyal (2000). A noncooperative model of network formation. *Econometrica 68*(5), 1181–1229.

[21] See chapter 7 of Skryms' Signals ([2010]) for an excellent discussion of this connection.

Barrett, J. A. (2020). Self-assembling games and the evolution of salience. *British Journal for the Philosophy of Science*.

Barrett, J. A., C. T. Cochran, S. Huttegger, and N. Fujiwara (2017). Hybrid learning in signalling games. *Journal of Experimental & Theoretical Artificial Intelligence 29*(5), 1119–1127.

Barrett, J. A. and B. Skyrms (2017). Self-assembling games. *The British Journal for the Philosophy of Science 68*(2), 329–353.

Barrett, J. A., B. Skyrms, and A. Mohseni (2019). Self-assembling networks. *The British Journal for the Philosophy of Science 70*(1), 301–325.

Beggs, A. W. (2005). On the convergence of reinforcement learning. *Journal of economic theory 122*(1), 1–36.

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

Herrmann, D. A. and B. Skyrms (forthcoming). Invention and evolution of correlated conventions. *The British Journal for the Philosophy of Science*.

Herrnstein, R. J. (1970). On the Law of Effect. *Journal of the Experimental Analysis of Behavior 13*(2), 243–266.

LaCroix, T. (2020a). On salience and signalling in sender–receiver games: partial pooling, learning, and focal points. *Synthese 197*(4), 1725–1747.

LaCroix, T. (2020b). Using logic to evolve more logic: Composing logical operators via self-assembly. *The British Journal for the Philosophy of Science*.

Lewis, D. (1969). *Convention: A philosophical study.* Harvard University Press.

Luce, R. D. (1959). *Individual choice behavior; a theoretical analysis.* New York: Wiley.

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* Oxford University Press.