Pyow-Hack

Order Sensitive Compositionality in Signaling Games

Nathan Gabriel

University of California Irvine January 29, 2021

Abstract

This paper presents a hierarchical Lewis-Skyrms signaling game that can develop a type of signaling system analogous to the alarm calls of putty-nosed monkeys (*Cercopithecus nictitans*). Putty-nosed monkeys display a compositional system of alarm calls with a semantics that is sensitive to the ordering of terms. This sensitivity to the ordering of terms has not been previously modeled with a Lewis-Skyrms signaling game literature. The hierarchical signaling game is paired with reinforcement (Roth-Erev) based learning dynamics. Simulation results show that reinforcement learning is sufficient for the development of a compositional signaling system analogous to putty-nosed monkey alarm calls.

1 Introduction

Compositionality is exhibited when a full statement in a language is comprised of component parts that contribute to its meaning. E.g. "the apple is red" and "the traffic light is red" are two distinct statements that have "red" as a component term. There are a number of ways of precisifying informal notions of compositionality. The types[1] of compositionality considered here are best understood in the context of the concrete models in which they obtain.

The literature on the evolution of compositionality in the context of signaling games includes a diversity of motivations. Franke [13] is inspired by Frege's emphasis of the role of compositionality in generating novel complex statements. Rather than being concerned with how compositionality could evolve, Steinert-Threlkeld [24] is interested in why our languages are compositional from a functional perspective. He presents a signaling game model in which a type of

[1] This paper utilizes four technical terms for making types of compositionality precise: sender-compositional, receiver-compositional, sender independent, and trivially compositional. To my knowledge, at least two of these terms make precise concepts that have not been directly discussed in the prior literature: receiver-compositional and sender independent. Sender-compositionality is a particularly weak type of compositionality that has been discussed, but not named, in the prior literature. Trivial compositionality has already been discussed and named in the literature.

negation plays a compositional role in signals and takes it to demonstrate the adaptive benefit of compositional signaling.

In a model that will be presented shortly, Barrett [7] is concerned with how a kind language might evolve and whether such kinds would be canonical. Section 2.1 outlines Barrett's model and discusses how it exhibits a relatively weak type of compositionality. Barrett, Cochran, and Skyrms [6] [5] introduce a hierarchical signaling game that can evolve a stronger type of compositionality. The hierarchical structure of the signaling game allows this compositionality to develop under simple reinforcement. In part, Barrett, Cochran, and Skyrms motivate their models with a gesture to a few simple types of compositionality that are exhibited by non-human animals.

The present paper is motivated by a type of compositionality that has been observed in putty-nosed monkey, *Cercopithecus nictitans*, alarm calls [22]. Using a hierarchical signaling game, the same type of game used by Barrett, Cochran, and Skyrms [6][5], this paper shows how a type of compositionality analogous to that of the putty-nosed monkey alarm calls can obtain. The extent to and way in which the models considered here are explanatory of putty-nosed monkeys' behavior is a topic beyond the purview of this paper.[2] Though, it might be noted that the simplest learning dynamic investigated, basic Roth-Erev reinforcement, bears some resemblance to the types of reinforcement learning exhibited by a wide array of species.[3] Putty-nosed monkeys display a compositional system of alarm calls with a semantics that is sensitive to the ordering of terms [22]. Specifically, this compositionality is such that there are component terms H and P along with statements HP and PH where the dispositions associated with HP and PH are different though reflective of the dispositions associated with H and P occurring in isolation. This is a type of compositionality that is new to the Lewis-Skyrms signaling game literature.[4]

After a brief description of putty-nosed monkey behavior, Section 2 presents some signaling game models already in the literature. Section 2.1 begins with a fairly rudimentary model and then proceeds to (Section 2.2) a description of a more complex hierarchical signaling game model in which basic reinforcement learning is supplemented with punishment via costly signals. The result of this is a gradual building up of the tools necessary to describe some models in which

[2]The philosophical terrain concerning the explanatory nature of various models is fairly thick. Buckner [9] and Weiskopf [28] examine models of cognition which they both hold to be explanatory for interventionist reasons, but they disagree with each other about whether those models are mechanistic. In contrast with those models, Chirimuuta [10] and Ross [21] examine models in neuroscience which prove to be explanatory precisely because they abstract away from details required for many types of interventions to be performed. Explicating where in this terrain lie the models from the pyow-hack game is a non-trivial task.

[3]Future research may explore the pyow-hack signaling game paired with various replicator dynamics rather than reinforcement based dynamics.

[4]Superficially, the Barrett, Cochran, and Skyrms hierarchical models seem to obtain this type of compositionality with statements 01 and 10 being associated with different dispositions (when a optimal signaling system obtains). However, this is superficial because the dispositions associated with a 0 in the first place of a compound statement are disjoint with the dispositions associated with a 0 in the second palace of a compound statement. This is discussed in more detail in at the end of Section 2.2.

a simplified version of the monkeys compositional calls can obtain. Section 3 presents the pyow-hack signaling game paired with a variety of reinforcement dynamics.

1.1 A Brief Description of Putty-nosed Monkey Behavior

Cercopithecus nictitans martini, putty-nosed monkeys, are a West African species. They typically live in groups of 13-22 individuals comprised of one adult male with several females and dependent juveniles [1]. Their common predators are crowned eagles and leopards. Group leaders give different alarm calls that correlate fairly robustly with the presence of leopards and eagles. They also have a call associated with group movement [1][2][4][22][23].

Putty-nosed monkey alarm calls are comprised of two basic calls: a hack (H) and a pyow (P). These basic calls are strung together in sequences of varying length. A sequence of repeated hacks, perhaps HHHHHHHH, is associated with aerial predators (eagles) and invokes the behavior of looking up. A sequence of repeated pyows, perhaps PPPPP, is associated with ground predators (leopards). Behaviorally, the pyow call sequences are associated with moving towards the caller since ground predators rely on stealth and the monkeys can collectively scare off the predator [4]. Sequences of pyows followed by hacks, perhaps PPPPHH, are associated with group movement. Sequences of hacks followed by pyows, perhaps HHHHPPP, occur when a nearby eagle moves away from the group. Longer call sequences seem to correlate with more urgent contexts when signaling for predators and increased distance traveled when signaling group movement; behaviorally, this correlates respectively with faster reaction times and potentially moving longer distances [3][4][22].

Schlenker et al. [22] give a detailed overview of putty-nosed monkey alarm calls, and reasons for interpreting the calls as semantically compositional. Additionally, they propose some possible referential or imperative semantics for the alarm calls.[5] In developing a compositional semantics, Schlenker et al. make a particularly insightful observation about the relation between calls associated with ground predators and calls associated with group movement. Though over a shorter distance, the monkeys move towards the caller when the call for a ground predator is issued (so they can collectively mob the predator). This provides some reason for interpreting a "pyow" as contributing similar meaning to the ground predator call as a "pyow" contributes to a group movement call.

The pyow-hack signaling game will simplify things by only allowing six different statements in the game: P, PP, PH, H, HH, and HP. On this simplification, putty-nosed monkey behavior translates to the following call system. When a leopard is nearby, the group leader issues a P call, to which group members are disposed to move towards the group leader. When a leopard is very nearby, the group leader issues a PP call, to which group members are disposed to quickly

[5]That is, the semantics is referential if PPPP means there's a leopard nearby and is imperative if PPPP is a command to move towards the caller. Schlenker et al. are incredibly thoughtful in their discussion of these semantics even if they do not think there is any strong reason to prefer one semantic over the other.

move towards the group leader. When moving, the group leader issues a PH call, to which group members are disposed to move an extended distance towards the group leader. When an eagle is nearby, the group leader issues a H call, to which group members are disposed to look up. When an eagle is very nearby, the group leader issues a HH call, to which group members are disposed to quickly look up. When a nearby eagle is leaving, the group leader issues a HP call, to which group members are disposed to look up and then elsewhere.

2 Compositionality in Signaling Games

This paper shows how a rudimentary form of compositionality, similar to the compositionality exhibited by putty-nosed monkeys, can be achieved through reinforcement learning in the context of hierarchical Lewis-Skyrms signaling game. This section begins with presenting a simple signaling game and gradually build in complexity to a hierarchical signaling game suitable to describing the basic communicative behavior of putty-nosed monkeys.

2.1 A Simple Signaling Game

A particularly simple form of compositionality that can be achieved with signaling games has a two-sender one-receiver structure and already exists in the established literature [7][8] [6] [5]. In this game there are four states of nature, four correspondingly appropriate actions, and two terms. This game is paired with basic Roth-Erev reinforcement.

The resulting model can be illustrated with a story. Suppose you and your partner host a weekly potluck. As guests arrive, you need to instruct a friend to bring either an appetizer or a dessert that is either hot or cold. Unfortunately, your friend's phone is broken and will only receive a single emoticon text from each of you containing either a smiley face, ☺, or a check mark, ✓. If you had planned for the broken phone, you could communicate the necessary information by having your partner send ☺ to indicate appetizer and ✓ for dessert, while you send ☺ to indicate hot and ✓ to indicate cold. Since this would lead to the appropriate action being performed for each of the four possible states, this is called a signaling system. However, for the sake of illustrating the learning dynamics, suppose further that each week you forget to establish such a protocol. Then, there is a simple reinforcement procedure you can each follow to establish such a protocol.

The reinforcement procedure is as follows. Take four buckets and apply one of the following labels to each of them: 'hot dessert', 'cold dessert', 'hot appetizer', and 'cold appetizer'. Now place two tennis balls in each bucket, one with a smiley face drawn on it and one with a check mark on it. Your partner does the same. Your friend with the broken phone labels four buckets with: '☺☺' , '☺✓', '✓☺', and '✓✓'. This corresponds to the four combinations of text messages that you and your partner can send. Your friend then puts four tennis balls in each bucket, each labeled with one of the four foods that might be brought: ‘hot dessert’, ‘cold dessert’, ‘hot appetizer’, and ‘cold appetizer’. Each week you and your partner randomly draw one ball from each of your buckets corresponding to the needed food. You and your partner each send the message indicated by the balls that you draw. Then, your friend randomly draws from the corresponding bucket to determine what food item she will bring. If your friend brings the correct item, you return the ball to the bucket and add an additional ball with the same label that you drew, else you return the ball without making an addition to the bucket’s contents.

The procedure followed on a single week is called a single play of the signaling game. Here’s an example of a successful play:

1. Nature chooses a state at random, with each state having equal probability of being selected. Suppose the state of requiring a hot dessert is chosen.
2. Observing that a hot dessert is needed. You go to the ‘hot dessert’ bucket and draw a ball at random with each ball having equal probability of being drawn. Suppose you draw a ☺.
3. You then text an emoji to your friend corresponding to the ball that was drawn; in this case a ☺.
4. Observing that a hot dessert is needed. Your partner goes to his ‘hot dessert’ bucket and draws a ball at random with each ball having equal probability of being drawn. Suppose he draws a ✓.
5. Your partner then texts an emoji to your friend corresponding to the ball that was drawn; in this case a ✓.
6. Your friend goes to the bucket corresponding to the signal that was transmitted, in this case ☺✓, and draws a ball at random with each ball having equal probability of being drawn. Suppose she draws a ball labeled ‘hot dessert’.
7. Corresponding to the ball that was drawn, your friend brings a hot dessert, perhaps a hot apple pie.
8. Since this action corresponds with the state of nature, this counts as a success.
9. Given the success, you replace the ball that you drew and add an additional ball with the same label as the one you drew to the bucket. In this case, you replace the ball you drew and add an additional ☺ ball to your ‘hot dessert’ bucket.
10. Given the success, your partner replaces the ball that he drew and adds an additional ball with the same label as the one he drew to the bucket. In this case, your partner replaces the ✓ ball that he drew and adds an additional ✓ ball to his ‘hot dessert’ bucket.
11. Given the success, your friend replaces the ball that she drew and adds an additional ball with the same label as the one she drew to the bucket. In this case, your friend replaces the ‘hot dessert’ ball that she drew and adds an additional ‘hot dessert’ ball to her ☺✓bucket.

12. In the event that your friend does not bring a food corresponding to the state of nature, this counts as a failure.

13. When a failure occurs, each person replaces the ball that was drawn and no additional balls are added to any of the buckets.

This concludes a single play of the game. A set of n repeated plays is called a n-run. The only way to know the outcome of a run is to actually perform all of the plays in a run. A strategy profile describes a sender or receiver's dispositions for all states of nature or signals that she can observe. Figure 1 shows a set of strategy profiles that could, in principle, be reached through the basic reinforcement procedure just described. When a set of strategy profiles describes dispositions such that for an state of nature the correct action is performed, it is called a signaling system. The set of strategy profiles in figure 1 describe a signaling system in this sense. A run is successful if it results in a signaling system.[6] The percentage of runs that result in a signaling system is the model's run success rate. In the game just described, with 10^6 plays per run under simple reinforcement learning, the run success rate is approximately 73% [7].

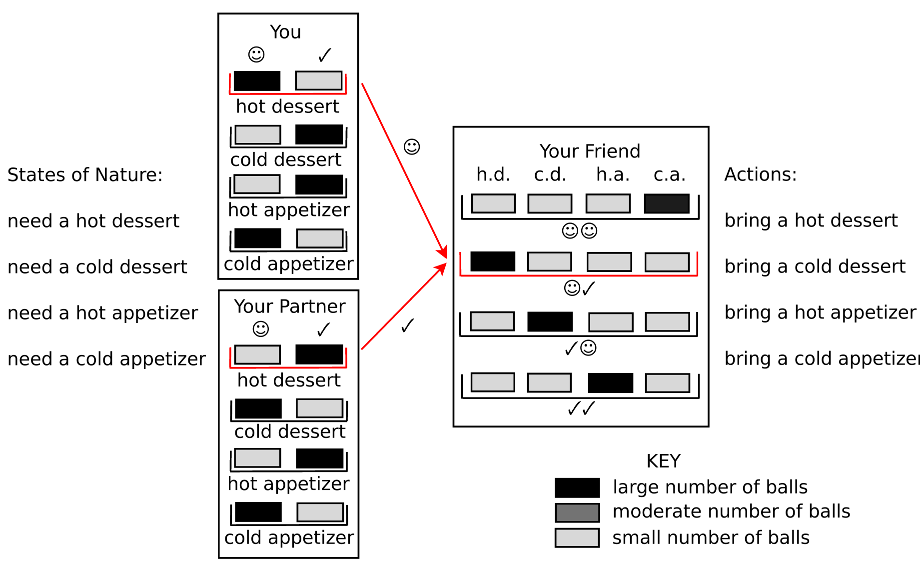

Figure 1: A Signaling System for the 4-state/2-sender/2-term Signaling Game. To indicate the quantities of different types of balls, each bucket is depicted with multiple boxes, one box for each type of ball in the bucket. The most likely balls to be drawn from a bucket are indicated the darker shaded in boxes.

[6] This is imprecise since the reinforcement dynamics generate mixed strategy profiles where, for any bucket, each ball type in the bucket has positive probability of being drawn. Consequently, true signaling systems are never reached, but they can be converged on. When the pyow-hack results are reported, a more precise definition of run success is used.

There is an important feature of the hypothetical run depicted Figure 1 worth noting. While the evolved meaning of the signals may agree with the meaning of terms in ordinary English, they need not. For example, in Figure 1, your partner's signals cohere with ordinary English in the sense that their sending '☺' indicates cold and '✓' indicates hot. However, your signals do not agree with ordinary semantics. When you send '☺', it indicates either a hot dessert or a cold appetizer since there are a large number of smiley faces in the hot dessert bucket and a large number of smiley faces in the cold appetizer bucket. It does not indicate hot because it is also sent in the event that a cold appetizer is required. Likewise, your '☺' cannot indicate our ordinary meaning associated with 'dessert' since it is also sent when a cold appetizer is required.

Another important aspect of the model just described is that separate senders in the model need not correspond to separate organisms. Suppose your partner is out of town, but also that your phone is now miraculously able to send two single character texts to your friend. You could realize a learning dynamics that is functionally identical to the original model by using your four buckets to dictate the first text that you send and your partner's buckets to determine the second text that you send. Separate senders in a model could correspond to separate functional mechanisms in an organism. There is evidence that different areas of the human brain realize different functional roles in language production and comprehension [12][16]. But, it is also possible that neurons within a particular brain area could be arranged to realize different functional roles.

2.1.1 A Brief Comparison Case

For the sake of comparison, consider a simpler signaling game. Suppose your partner is out of town. Rather than adopting your partner's buckets for a total of eight buckets, as described in the previous paragraph, in this game you use only four buckets. Each bucket is labeled with one of the four states of nature. For the initial state of the game, you place one each of four different types of ball corresponding to the four signals that you can send, '☺☺', '☺✓', '✓☺', and '✓✓'. Your friend's buckets are initialized exactly as they were in the previous game. On a play, a state of nature is determined at random and you draw from the corresponding bucket. You then send the signal indicated by your draw and your friend draws from the corresponding bucket to determine her action. As before learning occurs with simple reinforcement. When successes occur an additional ball of the drawn type is added to the bucket that was drawn from by both you and your friend. On failures, drawn balls are replaced with no additional balls added. Though only two emojis are utilized, this is in a four term game in which each pair of emojis is a unitary term. This is because each pair of emojis gets its own type of ball in your buckets. If instead of pairs of emojis, you labeled your four types of balls with 'w', 'x', 'y', and 'z' and your friend correspondingly changed her bucket labels, this would have no effect on the probability of a run resulting in an evolutionary stable equilibrium. The learning dynamics would remain the same.

Call this new four term game a 4-state/4-term game and the previously described two sender game a 4-state/2-term/2-sender game. The distinction between the 4-state/2-term/2-sender game and the 4-state/4-term game is substantial. Under simple reinforcement, runs of the 4-state/4-term game have a higher probability of success than runs of the 4-state/2-term/2-sender game [7]. Considering that both games can be understood as modeling different components of a single organism, we see that the differences between them reflect different possibilities for an organism's internal learning mechanisms.

2.2 Compositionality in a Hierarchical Signaling Game

There is a type of compositionality that can be exhibited in the 4-state/2-term/2-sender game. In the example from Figure 1, your "☺" is a term in the statement that is produced when a hot dessert is required. But it is also the case that it features in another composite statement, the statement transmitted when a cold appetizer is required. In this sense that "☺" is similar to "red" in the opening example. That is, "the apple is red" and "the traffic light is red" are two distinct statements that have "red" as a component term. Call a set of strategy profiles in a signaling game *sender-compositional* if there is a term that is transmitted as a component of at least two distinct statements.

Franke [13] expresses some dissatisfaction with compositionality that can be exhibited in the 4-state/2-term/2-sender game. Some of this dissatisfaction comes from the fact that it can only generate statements that are composite with respect to the senders' dispositions. The receiver acquires its dispositions by reinforcing actions as if each composite statement is unitary. Letting '0' and '1' be the two signals in the game, note that the receiver's reinforcement of an action picked from the 01 bucket has no direct effect on the contents of the 00 bucket or the 11 bucket. Call a game state *receiver-compositional* if there is a term that is a component of at least two distinct statements and reinforcing an action for one of the statements has a direct effect on the receivers dispositions towards the other statement(s) containing the given term.

Receiver-compositionality is intuitively desirable. Suppose someone learns the appropriate response to hearing "bring the red circle". If she simultaneously fails to learn the appropriate response to hearing "bring the red square", then it seems doubtful that she has learned the meaning of "red" as a component part of the statement "bring the red circle". Conversely, if learning the appropriate response to "bring the red circle" causes a person to be disposed to act appropriately in response to "bring the red square", then this seems more reflective of "red" being treated as a component term. This is what receiver-compositionality allows.

Barrett, Cochran and Skyrms [6][5] present a hierarchical signaling game that can produce game states that are both sender and receiver-compositional. The hierarchical signaling game has two basic senders, one executive sender, one basic receiver and one executive receiver. A state of nature features two binary properties and a context. For example, lets call the two properties temperature, which is either hot or cold, and dish, which is either appetizer or dessert. The context indicates what type of information is required for the receiver to perform the correct action on a given play. The game is sensitive to three contexts: temperature, dish, or both. This generates eight possible states of nature (dessert, appetizer, something hot, something cold, hot dessert, cold dessert, hot appetizer, cold appetizer).

In the game the executive sender sees the context and the basic senders see the properties. Correspondingly, the executive sender has three buckets and the basic senders have four buckets each, as they did in the previous game. The basic senders start the game with a 0 ball and a 1 ball in each of their buckets. The executive sender determines which of basic senders' signals gets transmitted. Her buckets contain three balls, a sender A ball, sender B ball and a both ball.

The basic receiver sees the transmitted signal and has four buckets: 00, 01, 10, 11. When only sender A's signal, 'x', is transmitted (x = 0 or 1), the basic receiver draws at random with equal probability from either the x0 or x1 bucket, with each ball in the given bucket having equal probability of being drawn. Similarly, the basic receiver draws at random with equal probability from either the 0x or 1x bucket when only sender B's signal, 'x', is transmitted. As in the previous game, each of the basic receiver's buckets has four types of balls in it. The executive receiver sees whether sender A, B or both transmitted a signal and determines whether the basic receiver's draw is interpreted as temperature, dish or both. Thus, the executive receiver has three buckets, A, B and Both, with three types of ball in each, temperature, dish or both.

Barrett, Cochran and Skyrms [6][5] add punishment to the learning dynamics through costly signals. On a play of the game there is a signal cost of 0.5 if only one of the basic senders' signals is transmitted and a signal cost of 1 if both basic senders transmit. Gross payoff is 0 if the incorrect action is performed, 1.5 if the correct action is performed and the context only requires sensitivity to a single property, and 2 if the correct action is performed and the context is complex. Net payoff is gross payoff minus the signal cost. On each play the buckets that were used in the play are reinforced with net payoff-many balls of the type that was drawn, with the addendum that no ball type is allowed to drop below 1. Note that signal-cost-many balls are being removed from buckets in the event of an incorrect action being performed. On this learning dynamic, with 10^8 plays per run, the run success rate is around 99.4% [6].[7]

[7] This is calculated with 0.98 cumulative success rate (average number of successes per play on a run) as the cutoff for a run being considered successful

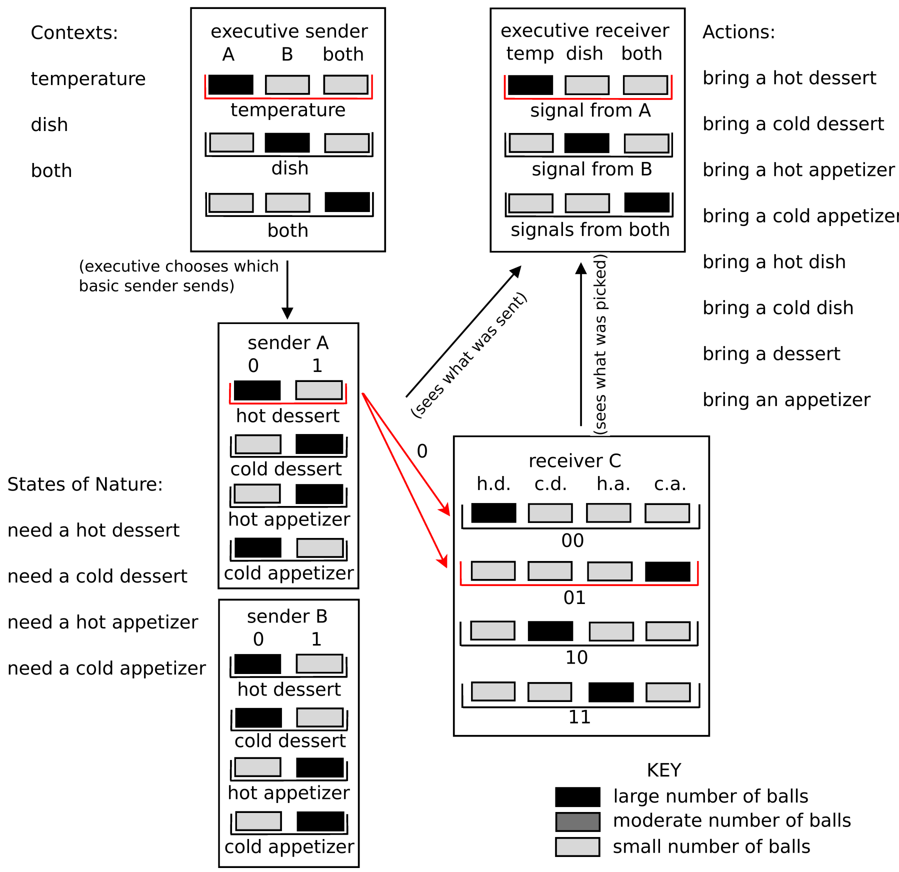

Figure 2: A Partial Pooling Equilibrium for the Hierarchical Signaling Game

For the purposes of this paper, call a set of strategy profiles a partial pooling equilibrium if they describe a suboptimal game state (i.e. a state that is not a signaling system) that can be expected to persist under the reinforcement learning dynamics.[8] Figure 2 portrays a partial pooling equilibrium that can be reached in this game. When only temperature is relevant for the correct action this set of strategy profiles performs at chance since sender A's signals cross cut both properties. Here sender A's dispositions are exactly the same as yours (the pot luck host's) in the example (Figure 1) of a successful solution to the 4-state/2-term/2-signaler game. The added sensitivity to simple and complex contexts makes sender A's cross cutting of properties no longer viable for a signaling system. However, sender A's terms can be thought of as having some meaning since sender A's dispositions contribute to successful communication

[8]This term is intentionally given a vague definition. For now, the state's being "expected to persist" can be read as sender and receiver's dispositions being expected to remain the same for a large number of plays, say 10^6 additional plays. Section 3.4 discusses Nash equilibria in the putty-nosed monkey model more precisely.

in complex contexts. Consider a hypothetical play of the game when this partial pooling equilibrium obtains:

1. Nature chooses a state and context at random with equal probability. Suppose the state of needing a hot dessert is chosen, but the context in which only temperature is relevant is chosen.
2. The executive sender observes the context and chooses a ball at random with equal probability from her temperature bucket. Suppose the executive sender chooses a sender A ball; this is the executive sender's most likely choice in the example equilibrium.
3. Since the sender A ball was chosen, only sender A will transmit a signal.
4. Sender A observes the state of needing a hot dessert. So, she draws at random with equal probability from her hot dessert bucket. Suppose she draws a 0 ball. Again, this is the most likely choice in the example equilibrium.
5. Given her draw, sender A transmits '0'.
6. Receiver C sees the '0' and it is determined at random with equal probability whether she will draw from the 00 bucket or 01 bucket. When only sender A transmits a signal and that signal is '0', the probability either the 00 or 01 bucket being determined remains 0.5 for the entirety of the game on every run of the game. Suppose it is determined that receiver C will draw from the 01 bucket.
7. Receiver C draws at random with equal probability from the 01 bucket. Suppose receiver C draws a cold appetizer ball. This is the ball that she is most likely to draw in the example equilibrium.
8. Receiver C's draw is now interpreted by the executive receiver.
9. The executive receiver sees that only sender A transmitted a signal and draws from her signal from sender A bucket. Suppose the executive receiver draws a temperature ball. In the example equilibrium, this is her most likely draw.
10. Given that the executive receiver drew a temperature ball, she interprets receiver C's draw, a cold appetizer ball, as needing a cold dish. So she brings a cold dish.
11. Since this action does not correspond with the state of nature, this counts as a failure.
12. Given the failure, each player returns the ball that she drew and no additional balls are added to the buckets that were drawn from.
13. Furthermore, given the failure, (signal cost)x(# of signals transmitted)-many balls of the type that was drawn are removed from the bucket that they were drawn from. In this case, the signal cost was 0.5 and only one signal was transmitted. So, the executive sender removes 0.5 sender A balls from her hot dessert bucket; sender A removes 0.5 0 balls from her

hot dessert bucket; receiver C removes 0.5 cold appetizer balls from her 01 bucket; and the executive receiver removes 0.5 temperature balls from her signal from sender A bucket.

14. When a success occurs, the drawn balls are returned and (gross payoff)-(signal cost)x(# of signals transmitted)-many balls of the drawn type are added to the buckets that they were drawn from.

This concludes a play of the game.

To see that the partial pooling equilibrium performs at chance when the state is hot dessert and the context is temperature only consider what would have happened if it had been determined that Receiver C would draw from the 00 bucket. In this case, Receiver C's most likely draw is a hot dessert ball and the executive receivers most likely draw is the temperature ball. These draws would lead to a success and 1.5 - (0.5)x(1) = 1 balls of the type drawn would be added to the buckets they were drawn from. This also helps explain why the example in Figure 2 is an equilibrium rather than being unstable. That is, it explains how all but receiver C's draws come to be reinforced. To see how receiver C drawing a cold appetizer ball from her 01 bucket gets reinforced, check that when the state is cold appetizer and the context is both each player drawing the most likely type of ball from their buckets leads to a success. Additionally, in the example equilibrium, receiver C will reinforce drawing a cold appetizer ball in about 50% of plays in which the state is either hot appetizer or cold appetizer and the context is dish only.

Though Figure 2 depicts a partial pooling equilibrium, it does exhibit both sender and receiver-compositionality. Sender-compositionality is exhibited as it was in the previous model. Receiver-compositionality is exhibited by sender B's terms. Assuming all of the senders and receivers pick the most populous balls in their buckets, then successfully signaling for a hot appetizer reinforces successfully signaling for an appetizer and vice versa. Barrett, Cochran and Skyrms's [6][5] introduction of executives to Lewis-Skyrms signaling games is what allows the possibility of receiver-compositionality obtaining. Specifically, when a basic receiver's draw can be given multiple interpretations by an executive, this allows the basic receiver's disposition to draw a particular ball from a particular bucket to be reinforced for multiple actions after an equilibrium has obtained. This is an alternative description of what is required to allow an equilibrium to obtain in which a term that is a component of at least two distinct statements and reinforcing an action for one of the statements has a direct effect on the receivers dispositions towards the other statement(s) containing the given term, the definition of receiver-compositionality.[9]

[9]The notion of receiver-compositionality could benefit from further refinement. Though the example partial pooling equilibrium performs at chance when the context is temperature, it is still the case in this context that about 50% of the time C's disposition to draw a hot dessert ball from the 00 bucket is reinforced. Since this disposition is also typically reinforced when the state is hot dessert and the context is both, this behavior seems to satisfy the current definition of receiver-compositionality. That is, there are two statements "0" and "00" that share the term "0" from sender A and receiver C reinforces the same disposition for both

Given this paper's focus on putty-nosed monkeys, there is a worry worth noting for this hierarchical signaling game. Consider interpreting the executive sender and basic senders as modeling different components of a single organism, and the receivers as modeling components of an organism distinct from the sender organism. Then, how can the transmission of a single '0' be interpreted? The game requires that the executive receiver is able to observe whether the '0' is transmitted by sender A or sender B. Perhaps the receiver organism is able to discriminate '0' transmitted by A from '0' transmitted by B in virtue of a tonal difference between two utterances. Such an interpretation implies that the game has four basic signals rather than two: 'A tone 0', 'B tone 0', 'A tone 1', and 'B tone 1'. Whether this hierarchical signaling game is interpreted as having two or four basic signals has no ramifications for attributing sender and receiver-compositionality to it. It still exhibits both. However, this worry highlights a way in which a game modeling putty-nosed monkey alarm calls must differ. The meaning of individual terms must be *sender independent.* A consequence of developing a game that allows this sender independence is that it allows the development of a novel form of compositionality.

3 The Pyow Hack Hierarchical Signaling Game

3.1 Structure of the Pyow-Hack Game

The pyow-hack signaling game abstracts and simplifies away from several of the details of putty nose monkeys' environment and behavior. Most noticeably, it only allows for call sequences of at most two signals. Like the Barrett, Cochran and Skyrms's model [6] [5], it is a hierarchical signaling game consisting of an executive sender, two basic senders, an executive receiver, and a basic receiver.

In the pyow-hack game there are six states of nature: a leopard is nearby, a leopard is very near (urgent), an eagle is nearby, an eagle is very near (urgent), a nearby eagle is moving away, and the group is moving. There are six corresponding appropriate actions: move towards caller, quickly move towards caller, look up, quickly look up, look up and elsewhere, and move an extended distance towards caller.

The executive sender as well as the basic senders can observe the state of nature. The executive sender determines whether just one or both of the basic senders will transmit a signal. This corresponds with the executive sender having six buckets, one for each state of nature. These buckets contain two types of balls, single transmission balls and dual transmission balls. As in the previous games, all of the players' buckets start with one ball of each type. The basic senders each have six buckets corresponding to the states of nature. The basic senders have two types of balls, P balls and H balls. On plays in which the executive sender draws a single transmission ball, it is determined at random

statements. This seems to clash with the type of compositionality that Franke is concerned with and could be remedied with the addition of a clause about the relevant statements leading to successful action at a rate above chance.

with equal probability whether sender A or sender B transmits a signal.[10]

The basic receiver has four buckets: PP, PH, HP, and HH. When as single P is transmitted, it is determined at random with equal probability whether the basic sender draws from the PP or PH bucket. When a single H is transmitted, it is determined at random with equal probability whether the basic sender draws from the HP or HH bucket. As in the previous hierarchical game, the basic receiver draws balls that can be given multiple interpretations by the executive. The basic receiver's buckets contain four types of balls labeled: (i) 'quickly move towards caller', (ii) 'move an extend distance towards caller', (iii) 'quickly look up', and (iv) 'look up and elsewhere'. These labels are the complex interpretations that the executive can give to the balls. Type (i) and (ii) balls can be given the simple interpretation "move towards the caller". Type (iii) and (iv) balls can be given the simple interpretation "look up". Thus, the executive receiver has two buckets, a single transmission bucket and a dual transmission bucket. Each of these buckets has two types of balls, simple interpretation balls and complex interpretation balls. This paper reports results from simulations of the pyow hack game for a variety of learning dynamics.

3.2 Simple Reinforcement Learning

Simple reinforcement learning (Roth-Erev) is the dynamic that was presented in Section 2.1. When a play is successful, drawn balls are returned to their buckets and one additional ball of the type drawn is added to the bucket that was drawn from for each player. On failures, balls are returned to the buckets they were drawn from. Here's an example play for the equilibrium depicted in Figure 3:

1. Nature chooses a state and context at random with equal probability. Suppose the state of a nearby leopard is chosen.
2. The executive sender observes the state and chooses a ball at random with equal probability from her nearby leopard bucket. Suppose the executive sender chooses a single transmission ball; this is the executive sender's most likely choice in the example equilibrium.
3. Since the single transmission ball was drawn, either sender A or sender B is chosen at random with equal probability to transmit a signal. Suppose sender A is chosen.
4. Sender A observes the state of a nearby leopard. So, she draws at random with equal probability from her leopard bucket. Suppose she draws a P ball. Again, this is the most likely choice in the example equilibrium.
5. Given her draw, sender A transmits 'P'.

[10] This is part of how the pyow hack game avoids the oddity discussed at the end of section two. In this game, executives are only sensitive to signal length. As will be seen shortly, the executive receiver cannot form dispositions relative to which basic sender transmitted a signal. She can only form dispositions relative to signal length.

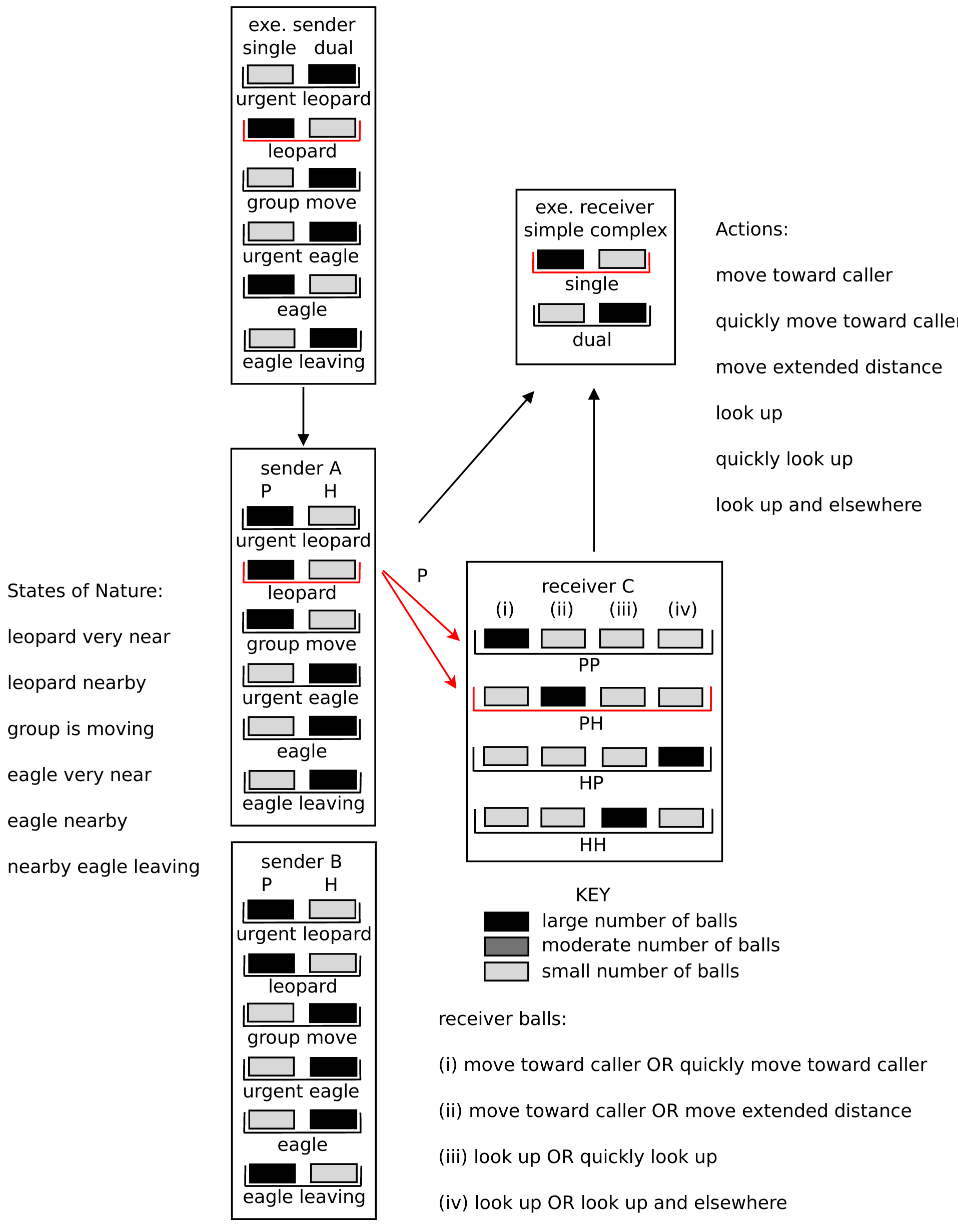

Figure 3: A Signaling System for the Pyow Hack Game

6. Receiver C sees the 'P' and it is determined at random with equal probability whether she will draw from the PP bucket or PH bucket. Suppose it is determined that receiver C will draw from the PH bucket.
7. Receiver C draws at random with equal probability from the 01 bucket. Suppose receiver C draws a (ii) ball. This is the ball that she is most likely to draw in the example equilibrium.
8. Receiver C's draw is now interpreted by the executive receiver.
9. The executive receiver sees that only single signal was transmitted and draws from her single bucket. Suppose the executive receiver draws a simple ball. In the example equilibrium, this is her most likely draw.
10. Given that the executive receiver drew a simple ball, she interprets receiver C's draw, (ii), as needing to move towards the caller. So this action is performed.
11. Since this is the correct action for the given state, this counts as a success.
12. Given the success, each player returns the ball that she drew along with an additional ball of the type that was drawn.
13. When a failure occurs, drawn balls are returned to the buckets that they were drawn from.

This concludes a play of the game.

When the pyow-hack game is played with basic reinforcement learning and 10^7 plays per run, the run success rate was 19.4% for 500 runs. This was calculated by means of measuring each run's cumulative success rate, the number of successful plays divided by the total number of plays. For 500 runs, 97 had a cumulative success rate above 0.92. This was an appropriate cutoff for determining whether a run was successful as $0.92 > 5.5/6$. That is, a cumulative success rate greater than 0.92 is indicative of plays being successful for each of the six states of nature. Several runs reached partial pooling equilibria with cumulative success rates around $0.83 \approx 5/6$ and $0.66 \approx 4/6$. This roughly corresponded with the equilibria generally being successful for five or four of the six states of nature.

Figure 4 depicts a partial pooling equilibrium generated from a run with a cumulative success rate of 0.825. This is a stable equilibrium in the sense that, for any given bucket, changing which type of ball is most populous will not improve the success rate.[11] For four of the six states, plays are successful when

[11] Equilibria are discussed more rigorously in Section 3.4. For the executive sender and executive receiver buckets, this is easy to see. Their buckets must be correlated with each other such that if the executive sender determines that a single signal is transmitted in simple contexts, then a the executive receiver's single bucket must have simple balls as most populous, and vice versa. It is also fairly easy to check this for the basic senders. E.g consider sender A's urgent leopard bucket. Most of the time H being the most populous ball leads to receiver C drawing from the HH bucket. Since (i) balls are most populous, this results in success. Supposing instead that P balls were the more populous balls in A's leopard bucket, then this would typically result in C drawing a (ii) ball from her PH bucket, and consequently failure.

Checking this fact for receiver C's buckets can be more involved (see footnote 11), though

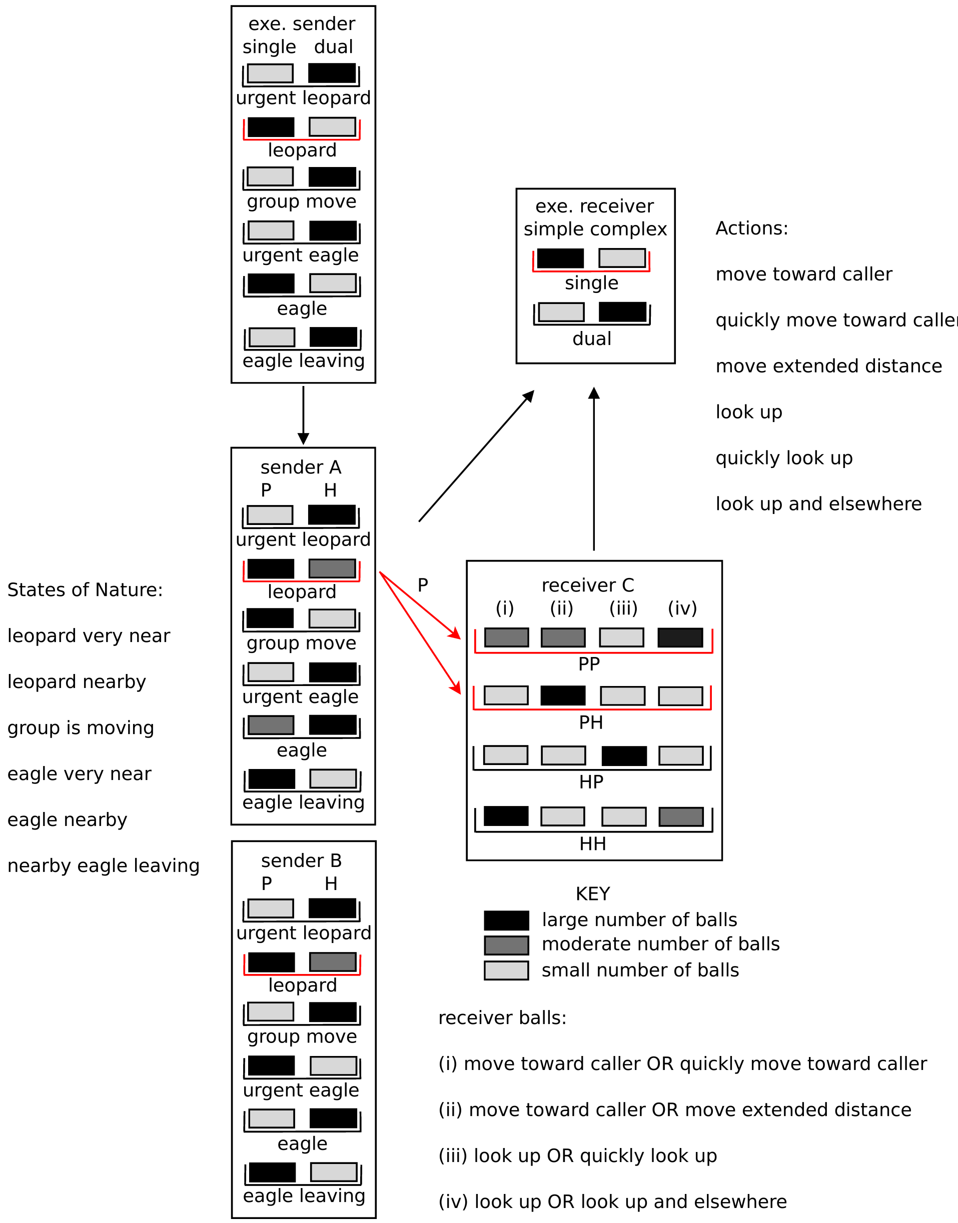

Figure 4: Partial Pooling Equilibrium with a Cumulative Success Rate of 0.825

each player picks the most likely ball from their bucket. However, this is not the case when the state of nature is either a nearby leopard or a nearby eagle. Consider what happens when the state of nature is a nearby leopard (marked in red). The executive sender will most likely draw a single-ball, meaning that at random with equal probability either sender A will transmit or sender B. Regardless of which basic sender transmits a signal, the most likely signal to be transmitted is a pyow. It is then determined at random with equal probability whether receiver C will draw from the PP or PH bucket. The executive receiver will draw from the single bucket and most likely give C's draw a simple interpretation. Thus, if C draws from the PH bucket the play will most likely be a success if the (ii)-ball is drawn, which is the most populous ball in the PH bucket. If C draws from the PP bucket, the play will most likely be a failure. So in general this equilibrium is successful about half of the time when the state is a nearby leopard. It behaves similarly when the state is a nearby eagle; it succeeds when C draws from the HP bucket and fails when C draws from the HH bucket.

Figure 5 depicts a the result of a run with a cumulative success rate of 0.666. This is a stable equilibrium in the sense that, for any given bucket, changing which type of ball is most populous will not improve the success rate.[12] Notice

---

it is not too difficult in this case. Consider the PP bucket. There are just two states that typically lead to C drawing from the PP bucket: leopard, and eagle leaving. In an eagle leaving state, C almost always draws from the PP bucket. In a leopard state C draws from the PP bucket about half the time, since a single P is transmitted and it is then determined at random with equal probability whether C draws from PP or PH. So it follow that, noting states occur at random with equal probability, about two thirds of C's draws from the PP bucket will be a result of an eagle leaving state. Pairing this with the fact that the (iv) ball is the only ball that can cause the appropriate action for a leaving eagle state, we get that the highest success rate, for plays in which the PP bucket is drawn from, occurs when (iv) balls are the most prevalent type of ball in the PP bucket.

It should be noted that this claim is made only with respect to the most prevalent type of ball in a given bucket. It is possible that decreasing the number of H balls in sender A's leopard bucket could marginally improve the success rate for occurrences of the leopard state (though one would expect this to also marginally lower the success rate in eagle leaving states; since this would cause (i) and (ii) balls to be more populous in the PP bucket).

[12]For the executive sender and executive receiver buckets, this is easy to see. Their buckets must be correlated with each other such that if the executive sender determines that a single signal is transmitted in simple contexts, then a the executive receiver's single bucket must have simple balls as most populous, and vice versa. It is also fairly easy to check this for the basic senders. E.g consider sender A's urgent leopard bucket. When a P ball is drawn, this almost always leads to success since receiver C's most populous ball type in the PP and PH buckets is the (i) ball (note that the executive typically draws a single ball from her urgent leopard bucket).

Checking this fact is a bit more involved for receiver C's buckets. The PP and PH buckets are fairly straightforward. Let's consider the HP bucket. There are three states that typically lead to receiver C drawing from the HP bucket: group movement, urgent eagle, and eagle leaving. Since each of these states occurs at random with equal probability, we expect each of these states to be responsible for C drawing from HP about a third of the time. (Note that this is because only a single signal is transmitted for each of these states. Thus about half of the time that one of these three states occur, C will draw from HP.) But now note that for each of these states, there is exactly one type of ball that can be drawn to lead to a successful action: (ii) for group movement, (iii) for urgent eagle, and (iv) for eagle leaving. So regardless of which of these three types of ball is most populous, we will expect success

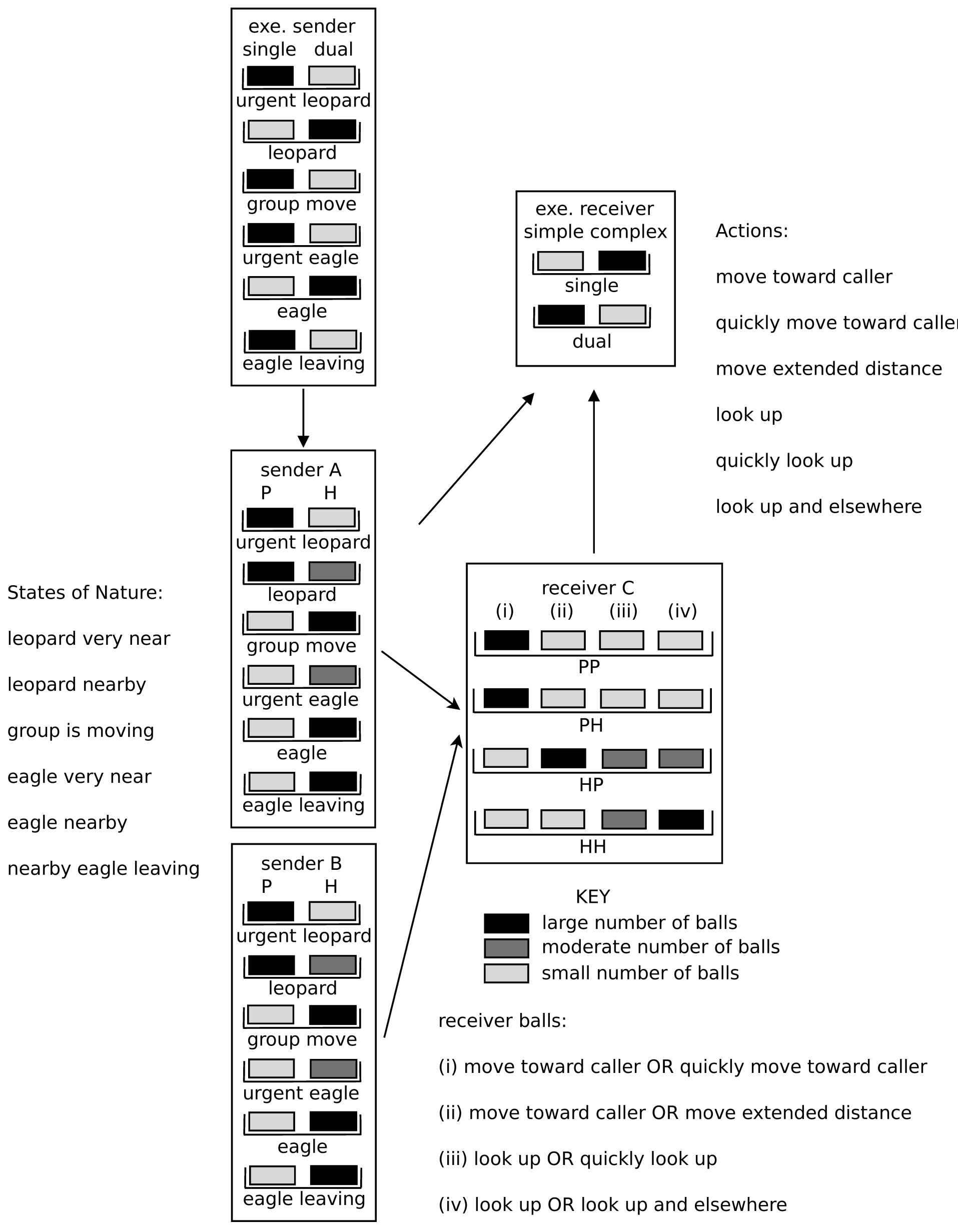

Figure 5: Partial Pooling Equilibrium with a Cumulative Success Rate of 0.666

that in this partial pooling equilibrium two signals are typically transmitted when the state of nature is simple, either nearby leopard or nearby eagle. For complex states, this equilibrium typically transmits a single signal. This contrasts with the previous partial pooling equilibrium with cumulative success of 0.825, in which single signals are transmitted in simple states and dual signals in complex states. This contrast was typical, but not universal, between partial pooling equilibria with cumulative success rates of 0.833 and 0.666. For signaling systems it is always the case that dual signals are transmitted in complex states. This is requisite since there are only two single signals that can be transmitted and four complex states.

3.3 Costly Signals and Partial Payoffs

Playing the game with reinforcement and punishment via costly signals substantially improves the run success rate. This was simulated with a gross payoff of 2 for success in any state, 0 for failures and a signal cost of 0.5 per term that was transmitted. As before: net payoff = gross payoff - signal cost. For example, consider the play depicted in figure 3. Since the play is a success gross payoff = 2 and since only one signal was transmitted signal cost = 0.5 yielding a net payoff of 1.5 balls of the type drawn being added to the buckets that they were drawn from. Failures result in 0.5 or 1 balls of the type drawn being removed from the buckets they were drawn from depending on whether one or two signals were transmitted. Simulating 500 runs with these payoffs and costs generated a run success rate of 57.4 %. As stated in figure 6, 31.2% of the runs resulted in partial pooling equilibria with cumulative success rates between 0.92 and 0.8; 9.6% of runs resulted in partial pooling equilibria with cumulative success rates between 0.8 and 0.6.

Now, the aim of this paper is to show how possibly putty-nosed monkeys could have acquired the alarm call system that they exhibit. Given this, consider the following worry. In the signaling game, if the state is a nearby leopard, and the receivers perform the action move quickly towards caller rather than merely moving towards the caller, then this counts as a failure. Under basic reinforcement the draws would not be reinforced; under reinforcement with costly signals the draws would be punished. However, this seems a bit implausible in

---

for approximately a third of C's draws from the HP bucket. The HH bucket can be similarly analyzed by noting that, while the eagle state will be responsible for a higher proportion of draws from the HH bucket, when compared to other states that lead to C drawing from HH, given the eagle state successful action will be performed regardless of whether C draws a (iii) or (iv) ball. That is, one can check that (iii) or (iv) balls being most populous results in the same probability of successful action when C draws from HH.

However, it is also the case that for some buckets the success rate can be marginally improved by decreasing the quantity of the second most populous type of ball. E.g consider sender A's leopard bucket. Decreasing the number of H balls will slightly improve the success rate. This is because, given the state of nature is a nearby leopard, the probability of failure when the receiver draws from the HP bucket is slightly higher than when she draws from the PP bucket. But most of the time drawing from the HP bucket, in the event that sender A transmits a H in the state of a nearby leopard, will lead to success. This is why sender A's leopard bucket has a moderate number of H balls.

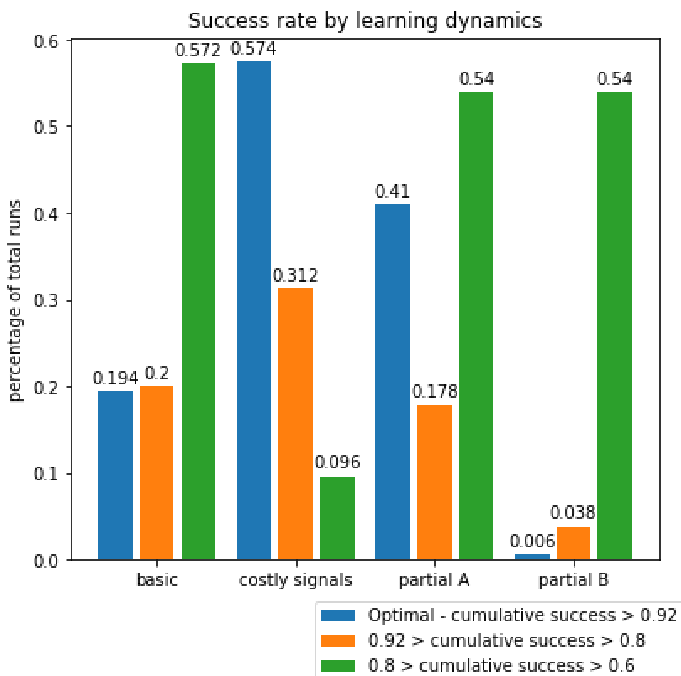

Figure 6: Success Rate be Learning Dynamics

the real world. Would a monkey not be equally rewarded for a more prompt performance of the same action? There are creative stories that can be told to mitigate this worry. Perhaps the increased energy expended for a fast response serves as a cost when such an expense is not requisite. Maybe the monkeys evaluate such plays as suboptimal.[13] Still, the worry is worth addressing more directly.

Consider the sets of similar states, $\mathrm{L_s}$ and $\mathrm{E_s}$, and similar actions, $\mathrm{L_a}$ and $\mathrm{E_a}$, letting: $\mathrm{L_s}$ ={leopard very near, leopard nearby, group is moving}, $\mathrm{L_a}$ = {mover toward caller, quickly move toward caller, move extended distance}, $\mathrm{E_s}$ = {eagle very near, eagle nearby, nearby eagle leaving}, and $\mathrm{E_a}$ = {look up, quickly look up, look up and elsewhere}. With these sets partial payoffs can be introduced. When the state is one of $\mathrm{L_s}$ and an incorrect action is performed, but that action is in $\mathrm{L_a}$, then gross payoff is the partial payoff value rather than zero. The same holds for a state in $\mathrm{E_s}$ and an incorrect action in $\mathrm{E_a}$. For example, suppose the partial payoff value is 0.5 and signal costs remain the same. Then, for a play with the state of a nearby leopard a single signal transmitted and the action move extended distance, the net payoff is (partial payoff)-(signal cost) = 0.5-0.5. So, zero additional balls of the type drawn would be added or removed from the buckets they were drawn from. For a partial payoff of 1, 0.5 additional balls of the type drawn would be added to the buckets they were drawn from. Figure 6 shows the results of 500 runs of the game with costly signals and partial payoff of 0.5 (partial A) and partial payoff of 1 (partial B). With a partial payoff of 0.5 the run success rate was 41%. A partial payoff of 1 gave a run success rate of 0.6%, though a majority of runs still reached equilibria that performed well above chance.

3.4 Equilibria Analysis

Until now equilibria have not been precisely defined. For any state of the pyow hack game, there are five strategy profiles, corresponding to the executive sender, sender A, sender B, receiver C, and the executive receiver. A mixed strategy profile is one in which an agent has probability p, $0 < p < 1$, of performing some output in response to a state of nature or a message. All strategy profiles generated from the described reinforcement learning are mixed (since buckets always contain at least one ball of each type). A strategy profile is pure if an agent determinantly produces a specific output for a given input. A *Nash equilibrium* is a game state in which none of the five agents can improve their average payoff by unilaterally changing her strategy profile.

Considering only pure strategies, there are 268435456 different combinations of strategy profiles in the pyow hack game. Of these, about 3% are Nash equilibria.[14] The mixed strategy profiles produced by reinforcement learning can be

[13] For this line of mitigation consider the justifications for actor-critic models of learning. Sutton and Bartol [26] give an introduction to actor critic models. For an example of an actor-critic model in the context of learning a behavior that is sensitive to the ordering of terms, see Hansen McKenzie and McClelland 2014 [14].

[14] 8027512 combinations are Nash equilibria.

transformed into pure strategy profiles by considering the most populous type of ball in each bucket as the only type of ball in the each respective bucket. After this transformation, nearly all game states produces by basic reinforcement learning are Nash equilibria.[15]

The story is more complicated when considering reinforcement supplemented with punishment via costly signals. Figure 7 depicts the issue. In this game state, the executive and basic senders all have exactly one ball of each type in their urgent leopard buckets. This is due to repeated punishment from the wrong action being performed when urgent leopard states occur. To see that the game state in figure 7 is not Nash, one can check that receiver C unilaterally changing her strategy profile to draw (i)-balls from her PP bucket (rather than (ii)-balls) would improve her average payoff. This type of game state, where senders have at least one bucket with exactly one ball of each type, occurs in almost all runs of reinforcement with punishment that do not produce signaling systems.

A natural reaction to the complication produced by reinforcement supplemented with punishment is to check what happens if the model alternates between reinforcement with punishment and basic reinforcement without punishment. Since the game state in Figure 7 is not Nash, it is not stable under basic reinforcement without punishment (though it seems to be stable under reinforcement with punishment). Therefore, one would expect the success rate of the state in Figure 7 to improve when the learning dynamic switches from reinforcement with punishment to reinforcement without punishment. Depending on how many iterations, comprised of first reinforcement with costly signals and then reinforcement without any punishment, are performed, the run success rate typically improve 7-11% over the run success rate of the costly signals model described in Section 3.2; i.e. run success rates ranged from 64-68%.[16] More sophisticated methods[17] of iterating different types of reinforcement and punishment generated runs success rates of up to 97.6%;[18] i.e. for certain parameters of the more sophisticated iterated reinforcement with punishment dynamics nearly every run generated a signaling system.[19] However, these more

---

[15] 93% of 10^6-runs are Nash and 96% of 10^7-runs are Nash.

[16] This was checked for number of iterations ranging from 2 to 15 and plays per run ranging between 10^7-runs and $6 * 10^7$-runs. Email ngabrie2@uci.edu for specific parameters and run success rates.

[17] These models introduced additional parameters such as: (i) imposing an upper bound on the number of balls of a given of a given type that could be in a bucket, and (ii) allowing a fixed base cost for failure on a given play (this could be used in conjunction with punishment via costly signals or independently as punishment absent costly signals.

[18] The run success rate of 97.6% was achieved with 8 iterations which summed to $3 * 10^7$ plays per run. Each iteration was equally composed of first (a)plays with gross payoff = 2, base punishment = 0 and signal cost of 0.5, followed by (b) plays with gross payoff = 5, base punishment = 0 and signal cost of 0.5, followed by (c) plays with gross payoff = 2, base punishment = 4 and signal cost of 0.5. Additionally the maximal number of balls of a given type in a given bucket was bounded by 10^5.

[19] This was checked for number of iterations ranging from 2 to 15 and plays per run ranging between 10^7-runs and $6 * 10^7$-runs. Email ngabrie2@uci.edu for specific parameters and run success rates.

sophisticated iterated reinforcement and punishment dynamics seem a bit ad hoc.[20]

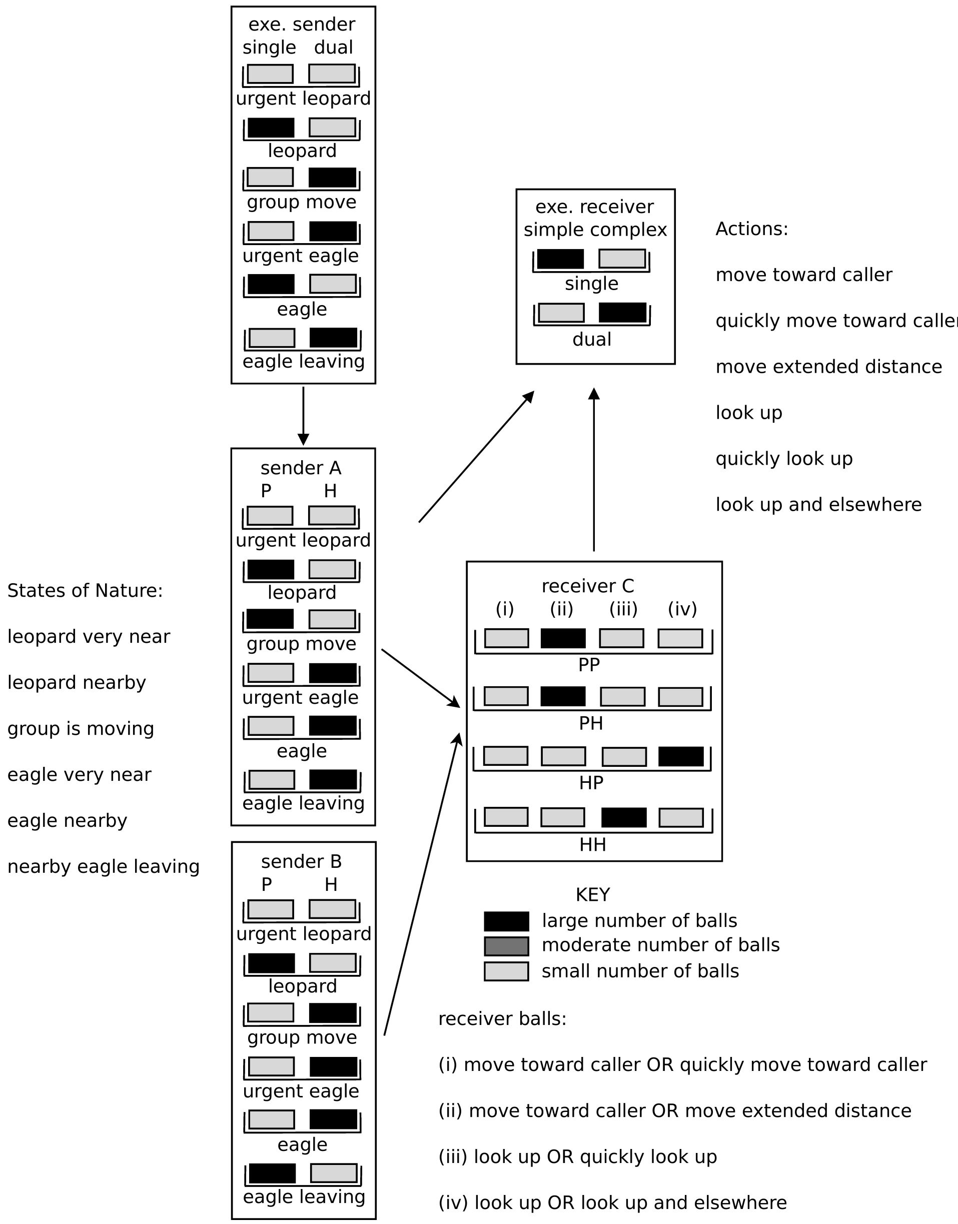

Figure 7: Stable Equilibrium that is not Nash

[20] This is particularly true if one is interested in developing something resembling an abstract mechanism sketch of how putty-nosed monkeys acquire a compositional system of alarm calls.

4 Discussion

4.1 Review of Technical Terms

Two types of compositionality have already been discussed: (i) a game state is *sender-compositional* if there is a term that is transmitted as a component of at least two distinct statements; (ii) a game state is *receiver-compositional* if there is a term that is a component of at least two distinct statements and reinforcing an action for one of the statements has a direct effect on the receivers dispositions towards the other statement(s) containing the given term. Receiver-compositionality allows a term to contribute similar dispositions to multiple statements that have the given term as a component part. Additionally, motivating some of the differences between the pyow-hack hierarchical game and the Barrett, Cochran, and Skyrms' [6][5] hierarchical game, a term in a stable game state is defined to be *sender independent* just in case transmission of the term typically results in the same action regardless of which sender transmitted it. The signaling game literature discusses a third type of compositionality defined by Schlenker et al. [23] and introduced to the signaling game literature by Steinert-Threlkeld [25][21]: (iii) a game state is *trivially compositional* just in case complex expressions are always interpreted by intersection (generalized conjunction) of the meanings of the parts of the expression. It can be checked that optimal signaling systems for the Barrett, Cochran, and Skyrms hierarchical game are trivially compositional.[22]

4.2 Order Sensitive Compositionality

Signaling systems for the pyow-hack game are not trivially compositional. If they were, since the terms are sender independent, HP would be associated with the same dispositions, as PH. But this cannot occur in a signaling system since the game only allows for six possible statements and requires distinct actions to be performed for each of the six states of nature. For a given game state, if transmitting HP and transmitting PH typically results in the same action being

[21] Although, it is not obvious that this is an appropriate definition for the sort of compositionality that Steinert-Threlkeld is concerned with. Steinert-Threlkeld constructs a game in which an artificial neural network is supposed to learn the function words "most" and "least" across different dimensions of properties that these function words can be applied to. In his current analysis, it is not clear that the network's learning the correct output for "most blue" contributes to its learning the correct output for "most green". But, this is the case for the human language that Steinert-Threlkeld is attempting to model. When a human learns to use "most" appropriately for some small domain of properties, she then is able to use the term appropriately for novel properties. This is because the content of the term is not specific to particular properties. Furthermore, there is no obvious impediment to Steinert-Threlkeld showing that his model exhibits the desired behavior. To do this, Steinert-Threlkeld should show that first training the network on one or two dimensions of properties allows it to learn "most" and "least" for a second or third dimension at a faster rate than it would with no pre-training. This is exactly the sort of generalization that artificial neural networks are valued for.

[22] Though it should be noted that the hierarchical game can exhibit non-trivial compositionality in suboptimal partial pooling equilibria [5].

performed in response to either statement, then at most only five of the six states of nature can be mapped to the correct action by the senders' and receiver's strategy profiles. This is a quick method of demonstrating that compositionality exhibited in the pyow-hack game is different from the compositionality exhibited in the Barrett, Cochran, and Skyrms hierarchical game. However, it does not show how compositionality with sensitivity to term ordering obtains in the pyow-hack game.[23]

Sensitivity to term ordering is allowed by the combination of both sender independent terms and receiver-compositionality. To see how this sensitivity is allowed, consider the signaling system diagrammed in Figure 3. It is easy to see that PH and HP are associated with different dispositions, actions. PH is typically transmitted when the state of nature is group movement and typically results in the action of moving an extended distance towards the caller. HP corresponds with the nearby eagle leaving state of nature and the look up & elsewhere action. However, this does not necessarily entail that the compositionality is sensitive to term ordering because of the worry described at the end of Section 2.2. Recall that this worry raises the concern that the P in PH is not the same term as the P in HP, perhaps one is an A-tone P and the other is a B-tone P making them functionally distinct terms. To establish that the P in the PH statement is the same term as the P in the HP statement, it must be shown that there is a connection between the dispositions associated with the P term in PH statements and the P term in HP statements. If one adopts a dispositional account of term meaning for the pyow-hack signaling game, then this amounts to a requirement that the P in PH and the P in HP have the same meaning.[24]

A clear dispositional connection between the P in PH statements and the second P in PP statements is as follows. Note that this is still following the signaling system depicted in Figure 3. The P in PH statements is connected to the P in solitary P statements by receiver-compositionality. In both the corresponding group movement and leopard states successful action reinforces the prevalence of (ii)-balls in receiver C's PH bucket.[25] Since terms are sender independent solitary P statements transmitted by sender A result in the same action as solitary P statements transmitted by sender B. That is, the reinforcement of P-balls in sender A's leopard bucket is connected to the reinforcement of P-balls in sender B's leopard bucket because when nearby leopard states occur the receiver probabilistically chooses from one of the same two buckets regardless of whether the solitary P is transmitted by sender A or B. The P in solitary P statements is connected to the second P in PP statements by receiver-compositionality. In both the corresponding leopard and urgent leopard states,

[23] This sensitivity to term order further highlights the differences between the pyow-hack game and the Barrett, Chochran, and Skyrms hierarchical game, which can exhibit non-trivial compositionality in suboptimal pooling equilibria, but cannot exhibit sensitivity to term order.

[24] Depending on how a dispositional account of meaning is understood, this might just translate to a requirement that the P in PH and P in HP have strongly related meanings.

[25] Note that (ii)-balls are only reinforced in the PH bucket 25% of the time that leopard states occur, since sender A transmits with 0.5 probability and receiver C then draws from the PH bucket with 0.5 probability.

successful action reinforces the prevalence of (i)-balls in receiver C's PP bucket. But this seems to be sufficient for establishing a dispositional connection between a P in PH statements and a P in HP statements. This is because the second P in a PP statement is transmitted by the same functional component, sender B, as the P in a HP statement; both P's being transmitted by sender B eliminates the worry raised at the end of section 2.2 that the two P's might in some way be different terms.

More abstractly, the pyow-hack game is able to realize a type of compositionality analogous to the compositionality exhibited by putty-nosed monkeys because of a combination of receiver-compositionality and sender independent terms. Receiver compositionality allows for a dispositional connection between a basic sender's term, say P, when used in a composite statement and when used in isolation. Sender independent terms allow for a dispositional connection between an isolated term transmitted by sender A and that same term transmitted by sender B. Thus there is an indirect connection between the dispositions associated with a term contributed by sender A to a composite statement and that same term contributed by sender B to a composite statement.

4.3 Final Remarks

Developing a Lewis-Skyrms signaling game that exhibits a compositionality analogous to the compositionality exhibited in putty-nosed monkey alarm calls generates a variety of interesting results. It shows how signaling games can be more directly connected to the behavior of a specific species. The pyow-hack signaling game motivates new terms for categorizing the types of compositionality that can be exhibited in signaling games, such as receiver-compositionality and sender independent terms. It highlights how a hierarchical signaling game can be used to develop a novel type of compositionality. In particular, Barrett, Chochran, and Skyrms' introduction of executive senders to signaling games makes the implementation of sender independent terms particularly easy.

This paper leaves as an open question the extent to which the pyow-hack game might be used to explain putty-nosed monkey behavior. If taken as providing a how-possibly explanation of putty-nosed monkey's acquisition of a compositional alarm call system, one might wonder whether the various models generate sufficiently high run success rates. This seems to deserve an affirmative answer. Having only basic reinforcement available to guide learning is a particularly hostile environment. Success rates from models with moderate partial payoffs show that small amounts of noise do not radically alter results. Furthermore, the pyow-hack signaling game abstracts away from a variety of real world details that might aid learning. In the real world, infant monkeys are born into an already established signaling system. This means that the established signaling system was likely selected over several generations allowing social groups with suboptimal signaling to die off. Additionally, in the real world visual cues can aid learning correct responses by reinforcement.

Some of the plausibility of taking the pyow-hack game to provide a how-possibly explanation of putty-nosed monkey alarm call behavior relies on it being plausible that the monkey's acquire their alarm call system through reinforcement learning. The non-exclusive alternative to this is that putty-nosed monkey's alarm call behavior is genetically hardwired. Thus, the pyow-hack game might motivate a number of research questions for comparative psychologists. For example, comparative psychologist might consider attempting to habituate putty-nosed monkeys to grammatical structures that they don't exhibit in their alarm calls. This could even be attempted with sounds not native to their environment.[26] E.g. experimenters could test whether they could condition a group to expect food when they hear a ABC pattern and expect a predator when they hear an ABA pattern. Success in such habituation would support the plausibility of the monkeys acquiring their signaling system via reinforcement learning.

The general simplicity of Lewis-Skyrms signaling games is a particularly beneficial feature of the pyow-hack game. As just stated, this can translate into showing how compositionality can arise in a particularly hostile environment. The game's simplicity also has the benefit of allowing it to be paired with a variety of different dynamics. Future research might explore how the game behaves when paired with various replicator dynamics, which might more immediately resemble evolutionary dynamics. So, the pyow-hack game suggests multiple paths to a putty-nosed monkey type of compositionality obtaining.

[26]Such an experiment would not be very dissimilar to the experiments that Neiworth et al. have had success with [17]

References

[1] Kate Arnold and Klaus Zuberbühler. The alarm-calling system of adult male putty-nosed monkeys, cercopithecus nictitans martini. *Animal Behaviour*, 72(3):643–653, 2006.

[2] Kate Arnold and Klaus Zuberbühler. Meaningful call combinations in a non-human primate. *Current Biology*, 18(5), 2008.

[3] Kate Arnold and Klaus Zuberbühler. Call combinations in monkeys: Compositional or idiomatic expressions? *Brain & Language*, page 303–309, 2012.

[4] Kate Arnold and Klaus Zuberbühler. Female putty-nosed monkeys use experimentally altered contextual information to disambiguate the cause of male alarm calls. *PLoS ONE*, 8(6), May 2013.

[5] J. A. Barrett, B. Skyrms, and C. Cochran. Hierarchical models for the evolution of compositional language. *IMBS Technical Report*, MBS(18), 2018.

[6] J. A. Barrett, B. Skyrms, and C. Cochran. On the evolution of compositional language. *Philosophy of Science*, forthcoming.

[7] Jeffrey A. Barrett. Dynamic partitioning and the conventionality of kinds*. *Philosophy of Science*, 74(4):527–546, 2007.

[8] Jeffrey A. Barrett. The evolution of coding in signaling games. *Theory and Decision*, 67(2):223–237, 2007.

[9] Cameron Buckner. Functional kinds: a skeptical look. *Synthese*, 192(12):3915–3942, 2014.

[10] M. Chirimuuta. Minimal models and canonical neural computations: the distinctness of computational explanation in neuroscience. *Synthese*, 191(2):127–153, 2013.

[11] Morten H. Christiansen and Nick Chater. *Creating language integrating evolution, acquisition, and processing*. The MIT Press, 2018.

[12] N. F. Dronkers, O. Plaisant, M. T. Iba-Zizen, and E. A. Cabanis. Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. *Brain*, 130(5):1432–1441, 04 2007.

[13] Michael Franke. The evolution of compositionality in signaling games. *Journal of Logic, Language and Information*, 25(3-4):355–377, 2015.

[14] Steven Stenberg Hansen, C. McKenzie, and J. McClelland. Two plus three is five: Discovering efficient addition strategies without metacognition. *Cognitive Science*, 36, 2014.

[15] Marc D. Hauser, Noam Chomsky, and W. Tecumseh Fitch. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598):1569–1579, Nov 2002.

[16] Margaret Naeser, Nancy Helm-Estabrooks, Gale Haas, Sanford Auerbach, and Malukote Srinivasan. Relationship between lesion extent in 'wernicke's area' on computed tomographic scan and predicting recovery of comprehension in wernicke's aphasia. *Archives of neurology*, 44:73–82, 02 1987.

[17] Julie J. Neiworth, Justin M. London, Michael J. Flynn, Deborah D. Rupert, Owen Alldritt, and Caleb Hyde. Artificial grammar learning in tamarins (saguinus oedipus) in varying stimulus contexts. *Journal of Comparative Psychology*, 131(2):128–138, 2017.

[18] Martin A. Nowak and David C. Krakauer. The evolution of language. *Proc. Natl. Acad. Sci. USA*, 96:8028–8033, Jul 1999.

[19] Mike Oaksford and Nick Chater. *Cognition and conditionals: probability and logic in human thinking*. Oxford University Press, 2012.

[20] U.S. Department of Health Office of the Assistant Secretary for Health and Human Services Office of the Secretary. Case summaries: Marc hauser, ph.d. harvard university. *Office of Research Integrity Newsletter*, 20(4):14–17, Sep 2012.

[21] Lauren N. Ross. Dynamical models and explanation in neuroscience. *Philosophy of Science*, 82(1):32–54, 2015.

[22] Philippe Schlenker, Emmanuel Chemla, Kate Arnold, and Klaus Zuberbühler. Pyow-hack revisited: Two analyses of putty-nosed monkey alarm calls. *Lingua*, 171:1–23, 2016.

[23] Philippe Schlenker, Emmanuel Chemla, Anne M. Schel, James Fuller, Jean-Pierre Gautier, Jeremy Kuhn, Dunja Veselinović, Kate Arnold, Cristiane Cäsar, Sumir Keenan, and et al. Formal monkey linguistics: The debate. *Theoretical Linguistics*, 42(1-2), 2016.

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

[25] Shane Steinert-Threlkeld. Towards the emergence of non-trivial compositionality. *Philosophy of Science*, forthcoming.

[26] Richard S. Sutton, Francis Bach, and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press Ltd, 2018.

[27] Daniel A. Weiskopf. The functional unity of special science kinds. *The British Journal for the Philosophy of Science*, 62(2):233–258, 2011.

[28] Daniel A. Weiskopf. Models and mechanisms in psychological explanation. *Synthese*, 183(3):313–338, 2011.