

# The Force of Innovation

## Emergence and Extinction of Messages in Signaling Games

Roland Mühlenbernd<sup>1</sup>, Jonas David Nick<sup>1</sup>, and Christian Adam<sup>1</sup>

University of Tübingen

**Abstract.** Lewis [L1] invented *signaling games* to show that meaning convention can arise simply from regularities in communicative behavior. The precondition for the emergence of such conventions are so-called *perfect signaling systems*. In a series of articles the emergence of such signaling systems was addressed by combining signaling games with learning dynamics; and not uncommonly researchers examined the circumstances aggravating the emergence of perfect signaling. It could be shown that especially by increasing the number of states, messages and actions for a signaling game perfect signaling becomes more and more improbable. This paper contributes to the question how the capability of innovation through emergence of new messages and extinction of unused messages would change these outcomes. Our results show that innovation in fact supports the emergence of perfect signaling.

## 1 Introduction

With the objective to explore the evolution of semantic meaning, signaling games recently became a leading model for this purpose. In line with this trend researchers used simulations to explore agents' behavior in repeated signaling games. Within this field of study two different lines of research are apparent: i) the simulation of a repeated 2-player signaling game combined with agent-based learning dynamics, in the majority of cases with the dynamics *reinforcement learning* (e.g. [B1], [BZ1], [S1]) and ii) evolutionary models by simulating population behavior, wherein signaling games are usually combined with the population-based *replicator dynamics* (e.g. [HH1], [HSRZ1]). To fill the gap between both accounts, recent work deals with applying repeated signaling games combined with agent-based dynamics on social network structures or at least multi-agents accounts (e.g. [Z1], [W1], [M1], [MF1]). With this paper we want to make a contribution to this line of research.

Barrett ([B1]) could show that i) for the simplest variant of a signaling game, called *Lewis game*, combined with a basic version of the learning dynamic *reinforcement learning* in a 2-player repeated game conventions of meaningful language use emerge in any case, but ii) by extending the domains<sup>1</sup> of the signaling game those conventions become more and more improbable. Furthermore the number of possible perfect signaling systems increases dramatically. This let surmise the motive that up to now researchers applied only the simple variant Lewis game on population and keep the hands off domain-extended signaling games. Because if even two players fail to learn

---

<sup>1</sup>With domains we refer to the sets of states, messages and action, which will be introduced in the following section.

perfect signaling from time to time, multiple players will not only have this problem, but also be confronted with an environment evolving to Babylon, where a great many of different signaling systems may evolve.

With this article we will show that by extending the learning dynamics to allow for innovation we can observe i) an improvement of the probability that perfect signaling emerges for domain-extended signaling games and ii) a restriction of the number of evolving perfect signaling systems in a population, even if the number of possible systems is huge. This article is divided in the following way: in Section 2 we'll introduce some basic notions of repeated signaling games, reinforcement learning dynamics and multi-agent accounts; in Section 3 we'll take a closer look at the variant of reinforcement dynamics we used - a derivative of Bush-Mosteller reinforcement; Section 4 is about how implementing innovation of new and extinction of unused messages significantly improves our results; we'll finish with some implications of our approach in Section 5.

## 2 Signaling Games and Learning

A signaling game  $SG = \langle \{S, R\}, T, M, A, Pr, U \rangle$  is a game played between a sender  $S$  and a receiver  $R$ . Initially, nature selects a state  $t \in T$  with prior probability  $\Pr(t) \in \Delta(T)$ <sup>2</sup>, which the sender observes, but the receiver doesn't.  $S$  then selects a message  $m \in M$ , and  $R$  responds with a choice of action  $a \in A$ . For each round of play, players receive utilities depending on the actual state  $t$  and the response action  $a$ . We will here be concerned with a variant of this game, where the number of states is on par with the number of actions ( $|T| = |A|$ ). For each state  $t \in T$  there is exactly one action  $a \in A$  that leads to successful communication. This is expressed by the utility function  $U(t_i, a_j) = 1$  if  $i = j$  and 0 otherwise. This utility function expresses the particular nature of a signaling game, namely that because successful communication doesn't depend on the used message, there is no predefined meaning of messages. A signaling game with  $n$  states and  $n$  messages is called an  $n \times n$ -game, whereby  $n$  is called the *domain* of the game.

### 2.1 Strategies and Signaling Systems

Although messages are initially meaningless in this game, meaningfulness arises from regularities in behavior. Behavior is defined in terms of strategies. A *behavioral sender strategy* is a function  $\sigma : T \rightarrow \Delta(M)$ , and a *behavioral receiver strategy* is a function  $\rho : M \rightarrow \Delta(A)$ . A behavioral strategy can be interpreted as a single agent's probabilistic choice or as a population average. For a  $2 \times 2$ -game, also called Lewis game, exactly two isomorphic strategy profiles constitute a perfect signaling system. In these, strategies are pure (i.e. action choices have probabilities 1 or 0) and messages associate states and actions uniquely, as depicted in Figure 1.

It is easy to show that for an  $n \times n$ -game the number of perfect signaling systems is  $n!$ . This means that while for a Lewis game we get the 2 signaling systems as mentioned above, for a  $3 \times 3$ -game we get 6, for a  $4 \times 4$ -game 24, and for a  $8 \times 8$ -game more than 40,000 perfect signaling systems. Moreover for  $n \times n$ -games with  $n > 2$  there is a possibility of partial *pooling equilibria*, which transmit information in a fraction of all possible cases.

---

<sup>2</sup> $\Delta(X) : X \rightarrow \mathbb{R}$  denotes a probability distribution over random variable  $X$ .



**Fig. 1.** Two perfect signaling systems of a  $2 \times 2$ -game, consisting of a pure sender and receiver strategy.

## 2.2 Models of Reinforcement Learning

The simplest model of reinforcement learning is *Roth-Erev reinforcement* (see [RE1]) and can be captured by a simple model based on urns, known as *Pólya urns*, which works in the following way: an urn contains balls of different types, each type corresponding to an action choice. Now drawing a ball means to perform the appropriate action. An action choice can be successful or unsuccessful and in the former case, the number of balls of the appropriate act will be increased by one, such that the probability for this action choice is increased for subsequent draws. All in all this model ensures that the probability of making a particular decision depends on the number of balls in the urn and therefore on the success of past action choices. This leads to the effect that the more successful an action choice is, the more probable it becomes to be elected in following draws.

But Roth-Erev reinforcement has the property that after a while the learning effect<sup>3</sup> slows down: while the number of additional balls for a successful action is a static number  $\alpha$ , in the general case  $\alpha = 1$ , as mentioned above, the overall number of balls in the urn is increasing over time. E.g. if the number of ball in the urn at time  $\tau$  is  $n$ , the number at a later time  $\tau + \epsilon$  must be  $m \geq n$ . Thus the learning effect is changing from  $\alpha/n$  to  $\alpha/m$  and therefore can only decrease over time.

*Bush-Mosteller reinforcement* (see [BM1]) is similar to Roth-Erev reinforcement, but without slowing the learning effect down. After a reinforcement the overall number of balls in an urn is adjusted to a fixed value  $c$ , while preserving the ratio of the different balls. Thus the number of balls in the urn at time  $\tau$  is  $c$  and the number at a later time  $\tau + \epsilon$  is  $c$  and consequently the learning effect stays stable over time at  $\alpha/c$ .

A further modification is the adaption of *negative reinforcement*: while in the standard account unsuccessful actions have no effect on the urn value, with negative reinforcement unsuccessful communication is punished by decreasing the number of balls leading to an unsuccessful action.

By combining Bush-Mosteller reinforcement with negative reinforcement, the resulting learning dynamic follows the concept of *lateral inhibition*. In particular, a successful action will not only increase its probability, but also decrease the probability of competing actions. In our account lateral inhibition applies to negative reinforcement as well: for an unsuccessful action the number of the appropriate balls will be decreased, while the number of each other type of ball will be increased.

---

<sup>3</sup>The learning effect is the ratio of additional balls for a successful action choice to the overall number of balls.

### 2.3 Applying Reinforcement Learning on Repeated Signaling Games

To apply reinforcement learning to signaling games, sender and receiver both have urns for different states and messages and make their decision by drawing a ball from the appropriate urn. We assume the states are equally distributed. The sender has an urn  $\mathcal{U}_t$  for each state  $t \in T$ , which contains balls for different messages  $m \in M$ . The number of balls of type  $m$  in urn  $\mathcal{U}_t$  designated with  $m(\mathcal{U}_t)$ , the overall number of balls in urn  $\mathcal{U}_t$  with  $|\mathcal{U}_t|$ . If the sender is faced with a state  $t$  she draws a ball from urn  $\mathcal{U}_t$  and sends message  $m$ , if the ball is of type  $m$ . Accordingly the receiver has an urn  $\mathcal{U}_m$  for each message  $m \in M$ , which contains balls for different actions  $a \in A$ , whereby the number of balls of type  $a$  in urn  $\mathcal{U}_m$  designated with  $a(\mathcal{U}_m)$ , the overall number of balls in urn  $\mathcal{U}_m$  with  $|\mathcal{U}_m|$ . For a receiver message  $m$  the receiver draws a ball from urn  $\mathcal{U}_m$  and plays the action  $a$ , if the ball is of type  $a$ . Thus the sender's behavioral strategy  $\sigma$  and receiver's behavioral strategy  $\rho$  can be defined in the following way:

$$\sigma(m|t) = \frac{m(\mathcal{U}_t)}{|\mathcal{U}_t|} \quad (1) \quad \rho(a|m) = \frac{a(\mathcal{U}_m)}{|\mathcal{U}_m|} \quad (2)$$

The learning dynamics are realized by changing the urn content dependent on the *communicative success*. For a Roth-Erev reinforcement account with a positive update value  $\alpha \in \mathbb{N} > 0$  and a lateral inhibition value  $\gamma \in \mathbb{N} \geq 0$  the following update process is executed after each round of play: if communication via  $t$ ,  $m$  and  $a$  is successful, the number of balls in the sender's urn  $\mathcal{U}_t$  is increased by  $\alpha$  balls of type  $m$  and reduced by  $\gamma$  balls of type  $m' \neq m$ . Similarly, the number of balls in the receiver's urn  $\mathcal{U}_m$  is increased by  $\alpha$  balls of type  $a$  and reduced by  $\gamma$  balls of type  $a' \neq a$ .

Furthermore for an account with negative reinforcement urn contents also change in the case of unsuccessful communication for the negative update value  $\beta \in \mathbb{N} \geq 0$  in the following way: if communication via  $t$ ,  $m$  and  $a$  is unsuccessful, the number of balls in the sender's urn  $\mathcal{U}_t$  is decreased by  $\beta$  balls of type  $m$  and increased by  $\gamma$  balls of type  $m' \neq m$ ; the number of balls in the receiver's urn  $\mathcal{U}_m$  is decreased by  $\beta$  balls of type  $a$  and increased by  $\gamma$  balls of type  $a' \neq a$ . The lateral inhibition value  $\gamma$  ensures that the probability of an action can get zero and it speeds up the learning process.

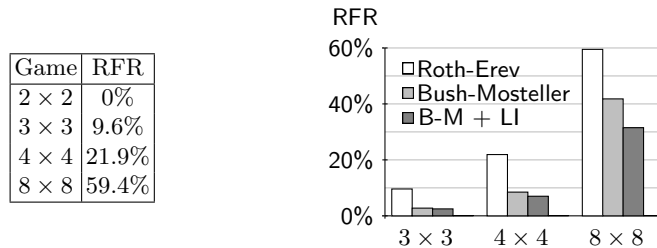
We extended the Bush-Mosteller reinforcement for applying it to games with more than two messages. The content of the appropriate sender and receiver urns will be adjusted to a predefined value in the following way: for the given value  $c$  of fixed urn content it is assumed that before a round of play the urn content of all sender and receiver urns  $|\mathcal{U}| = c$ . After a round of play it may be the case that the urn content  $|\mathcal{U}| = d \neq c$ . Now the number  $n_i$  of each type of ball  $i$  is multiplied by  $c/d$ .<sup>4</sup> For two messages the Bush-Mosteller is equivalent to our extension by setting the learning parameter of the original model to  $\phi = \frac{c-\alpha}{c+\alpha}$ .

### 2.4 Multi-Agent Accounts

It is interesting not only to examine the classical 2-players sender-receiver game, but the behavior of agents in a society (e.g. [Z1], [W1], [M1], [MF1]): more than 2 agents interact with each other and switch between sender and receiver role. In this way an agent can learn a sender and a receiver strategy as well. Now if such a combination forms a signaling system, it is called a *signaling language* and the corresponding agent

---

<sup>4</sup>In this account urn contents and numbers of balls are real numbers



**Fig. 2.** *Left:* Barrett’s results for different  $n \times n$  games. *Right:* Comparison of different learning dynamics: Barrett’s results of Roth-Erev reinforcement, results for Bush-Mosteller reinforcement without and with lateral inhibition.

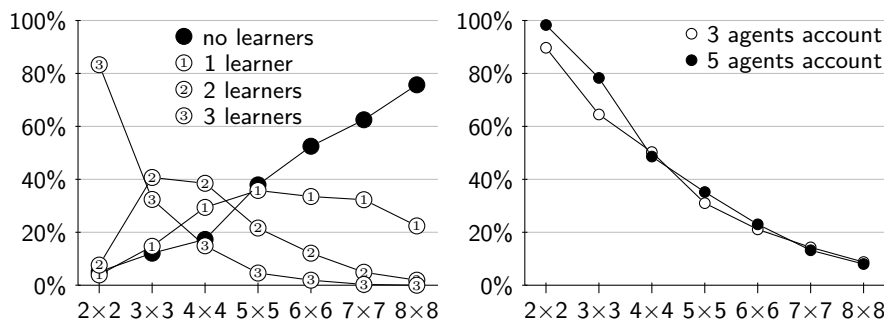
is called a *learner*. Thus the number of different possible signaling languages is defined by the number of possible signaling systems and therefore for a  $n \times n$ -game there are  $n!$  different languages an agent can learn. Furthermore if an agent’s combination of sender and receiver strategy forms a pooling system, it is called a *pooling language*. After all it is easy to show that the number of possible pooling languages outvalues the number of possible signaling languages for any kind of  $n \times n$ -game.

### 3 Simulating Bush-Mosteller

Barrett (see [B1]) simulated repeated signaling games with Roth-Erev reinforcement in the classical sender-receiver variant and computed the *run failure rate* (RFR). The RFR is the proportion of runs not ending with communication via a perfect signaling system. Barrett started  $10^5$  runs for  $n \times n$ -games with  $n \in \{2, 3, 4, 8\}$ . His results show that 100% (RFR = 0) of  $2 \times 2$ -games were successful. But for  $n \times n$ -games with  $n > 2$ , the RFR increases rapidly (Figure 2, left).

To compare different dynamics, we started two lines of simulation runs for Bush-Mosteller reinforcement in the sender-receiver variant with urn content parameter  $c = 20$  and reinforcement value  $\alpha = 1$ . For the second line we additionally used lateral inhibition with value  $\gamma = 1/|T|$ . We tested the same games like Barrett and correspondingly  $10^5$  runs per game. In comparison with Barrett’s findings our simulation outcomes i) resulted also in a RFR of 0 for the  $2 \times 2$ -game, but ii) revealed an improvement with Bush-Mosteller reinforcement for the other games, especially in combination with lateral inhibition (see Figure 2, right). Nevertheless, the RFR is never 0 for  $n \times n$ -games with  $n > 2$  and gets worse for increasing  $n$ -values, independent of the dynamics.

To analyze the behavior of agents in a multi-agent account, we started with the smallest group of agents in our simulations: three agents arranged in a complete network. In contrast to our first simulations all agents communicate as sender and as receiver as well and can learn not only a perfect signaling system, but a signaling language. Furthermore it was not only to examine if the agents have learned a language, but how many agents learned one. With this account we started between 500 and 1000 simulation runs with Bush-Mosteller reinforcement ( $\alpha = 1$ ,  $c = 20$ ) for  $n \times n$ -games with  $n = 2 \dots 8$ . Each simulation run stopped, when each agent in the network has learned a signaling language or a pooling language. We measured the percentage of simulation runs ending with no, one, two or three agents, which have learned a signaling language.



**Fig. 3.** Left: Percentage of simulation runs ending with a specific number of learners on signaling languages in a network with three agents for different  $n \times n$ -games with  $n = 1 \dots 8$ . Right: Average percentage of agents learning a signaling language over all runs for different  $n \times n$ -games with  $n = 1 \dots 8$ . Comparison of the results of a complete network of 3 agents (white circles) and 5 agents (black circles).

We got the following results: for a  $2 \times 2$ -game, all three agents have learned the same signaling language in more than 80% of all simulation runs. But for a  $3 \times 3$ -game in less than a third of all runs agents have learned a signaling language; in more than 40% of all runs two agents have learned a signaling language and the third one a pooling language. And it gets even worse for higher  $n \times n$ -games. E.g. for an  $8 \times 8$ -game in almost 80% of all runs no agents have learned a signaling language and never have all agents learned a signaling language. Figure 3 (left) depicts the distribution of how many agents have learned a signaling language (no learner, only one learner, two learners or all three agents are learners of a signaling language) for  $n \times n$ -games for  $n = 2 \dots 8$ .<sup>5</sup>

In addition we were interested in whether and how the results would change by extending the number of agents. Thus in another line of experiments we tested the behavior of a complete network of 5 agents for comparison with the results of the 3 agents account. Figure 3 (right) shows the average number of agents who learned a signaling language per run for different  $n \times n$ -games. As you can see for  $2 \times 2$ -games and  $3 \times 3$ -games the enhancement of population size leads to a higher average percentage of agents learning a signaling language. But for games with larger domains the results are by and large the same.

The results for the classical sender-receiver game reveal that by extending learning accounts the probability of the emergence of perfect signaling systems can be improved but nevertheless is never one for an  $n \times n$ -game, if  $n$  is large enough. Furthermore the results for the multi-agent account with only three agents show that even for a  $2 \times 2$ -game not in any case all agents learn a language. And for games with larger domains, results get worse. Furthermore results don't get better or worse by changing the number of agents, as shown in a multi-agent account with 5 agents. But how could natural languages arise by assuming them having emerged from  $n \times n$ -games with a

<sup>5</sup>Note: further tests with Bush-Mosteller reinforcement in combination with negative reinforcement and/or lateral inhibition revealed that in the same cases the results could be improved for  $2 \times 2$ -games, but were in any case worse for all other games with larger domains.

huge  $n$ -value and in a society of much more interlocutors? We'll show that by allowing for the extinction of unused messages and the emergence of new messages, perfect signaling systems emerge for huge  $n$ -values and multiple agents in any case. In other words, we'll show that stabilization needs innovation.

## 4 Innovation

The idea of innovation in our account is that messages can become extinct and new messages can emerge, thus the number of messages during a repeated play can vary, whereas the number of states is fixed. The idea of innovation and extinction for reinforcement learning applied on signaling games stems from Skyrms (2010), whereby to our knowledge it is completely new i) to combine it with Bush-Mosteller reinforcement plus negative reinforcement and ii) to use it for multi-agent accounts.

The process of the emergence of new messages works like this: additionally to the balls for each message type each sender urn has an amount of *innovative balls* (according to Skyrms we call them *black balls*). If drawing a black ball the sender sends a completely new message, not ever used by any agent of the population. Because the receiver has no receiver urn of the new message, he chooses a random action. If action and state matches, the new message is adopted in the *set of known messages* of both interlocutors in the following way: i) both agents get a receiver urn for the new message, wherein the balls for all actions are equiprobable distributed, ii) both agents' sender urns are filled with a predefined amount of balls of the new message and iii) the sender and receiver urn involved in this round are updated according to the learning dynamic. If the newly invented message doesn't lead to successful communication, the message will be discarded and there will be no change in the agents strategies.

As mentioned before, messages can become extinct, and that happens in the following way: because of lateral inhibition, infrequently used or unused messages' value of balls in the sender urns will get lower and lower. At a point when the number of balls of a message is 0 for all sender urns, the message isn't existent in the active use of the agent (i.o.w. she cannot send the message anymore), and will also be removed from the agent's passive use by deleting the appropriate receiver urn. At this point the message isn't in this agent's set of known messages. Besides, there is no other interference between sender and receiver urn of one agent.

Some further notes:

- it is possible that an agent can receive a message that is not in her set of known messages. In this case she adopts the new message like described for the case of innovation. Note that in a multi-agent setup this allows for a spread of new messages
- the black balls are also affected by lateral inhibition. That means that the number of black balls can decrease and increase during runtime; it can especially be zero
- a game with innovation has a dynamic number of messages starting with 0 messages, but ends with  $|M| = |T|$ . Thus we call an innovation game with  $n$  states and  $n$  ultimate messages an  $n \times n^*$ -game

### 4.1 The Force of Innovation

The total number of black balls of an agent's sender urns describes his personal *force of innovation*. Note that black balls can only increase by lateral inhibition in the case

Game	$2 \times 2^*$	$3 \times 3^*$	$4 \times 4^*$	$5 \times 5^*$	$6 \times 6^*$	$7 \times 7^*$	$8 \times 8^*$
3 agents	1,052	2,120	4,064	9,640	21,712	136,110	> 500,000
5 agents	2,093	5,080	18,053	192,840	> 500,000	> 500,000	> 500,000

**Table 1.** Runtime Table for  $n \times n^*$ -games with  $n = 2 \dots 8$ ; for a complete network of 3 agents and 5 agents.

of unsuccessful communication and decrease by lateral inhibition in the case of successful communication. This interrelationship leads to the following dynamics: successful communication lowers the personal force of innovation, whereas unsuccessful communication raises the personal force of innovation. If we define the global *force of innovation* for a group of connected agents  $X$  as the average personal force of innovation over all  $x \in X$ , then the following holds: the better the communication between agents in a group  $X$ , the lower the global force of innovation of this group and vice versa. In other words, this account realizes a plausible social dynamics: if communication works, then there is no need to change and therefore a low (or zero) value of the force of innovation, whereby if communication doesn't work, the force of innovation rises.

## 4.2 Learning Languages by Innovation: A Question of Time

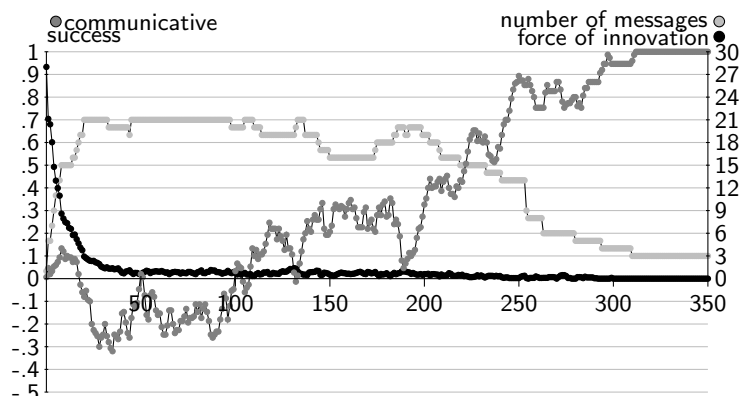
We could show in section 3 that the percentage of agents learning a signaling language in a multi-agent context is being decreased by increasing the domain size of the game. To find out whether innovation can improve these results we started simulation runs with the following settings:

- network types: complete network with 3 agents and with 5 agents
- learning dynamics: Bush-Mosteller reinforcement with negative reinforcement and lateral inhibition value ( $\alpha = 1$ ,  $\beta = 1$ ,  $\gamma = 1/|T|$ ) and innovation
- initial state: every urn of the sender is filled with black balls and the receiver does not have any a priori urn.
- experiments: 100 simulation runs per  $n \times n^*$ -game with  $n = 2 \dots 8$
- break condition: simulation stops if the communicative success of every agents exceeds 99% or the runtime passes the runtime limit of 500,000 communication steps (= runtime)

These simulation runs gave the following results: i) for the 3-agents account in combination with  $n \times n^*$ -games for  $n = 2 \dots 7$  and the 5-agents in combination with  $n \times n^*$ -games for  $n = 2 \dots 5$  all agents have learned a signaling language in each simulation run and ii) for the remaining account-game combinations all simulation runs exceeded the runtime limit (see Table 1). We expect that for the remaining combination all agents will learn a signaling language as well, but it takes extremely long.

All in all we could show that the integration of innovation and extinction of messages leads to a final situation where all agents have learned the same signaling language, if the runtime doesn't exceed the limit. Nevertheless we expect the same result for account-game combinations where simulations steps of these runs exceeded our limit for a manageable runtime.





**Fig. 4.** Simulation run of a  $3 \times 3^*$ -game with innovation in a 3-agents population. Communicative success, number of used messages and force of innovation of all agents in the population; number of simulation steps at  $x$ -axis.

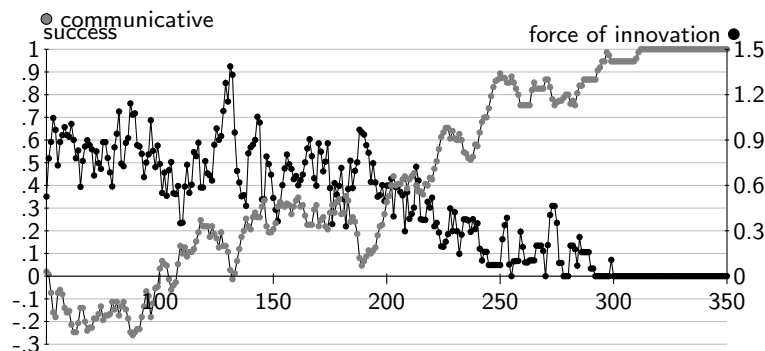
### 4.3 The Development of Signaling Languages by Innovation

As our experiments in the last section showed, by applying Bush-Mosteller reinforcement learning with innovation all agents learn the same signaling language for a small group of agents and any  $n \times n^*$ -game with  $n = 2 \dots 7$ . Let's take a closer look at how a  $3 \times 3^*$ -game develops during a simulation run by analyzing i) one randomly chosen agent's parameters and ii) parameters of the whole population. Three parameters are of interest to us:

- *communication success*: utility value averaged over the last 20 communication steps averaged over all agents in the population
- *number of messages in use*: number of actually used messages in the whole population
- *force of innovation*: absolute number of black balls averaged over all agents

Figure 4 shows the resulting values for the whole population: in the beginning all the agents try out a lot of messages, which reduces the number of black balls in the urns because balls for the new messages are added and then the urn content is normalized. Note that for the first communication steps the force of innovation drops rapidly, while the number of messages rises until it reaches 21 messages here. As you can see in the course of the success-graph, the work is not done here. Once they have more or less agreed on which messages might be useful, the agents are trying them out and it is only when finally a subset of those messages is probabilistically favored that the success is increasing, while the number of known messages decreases, until the success finally reaches a perfect 1 on average, while the number of messages equals that of the states (3) and the force of innovation is zero.

What you can see in the figures as well is that even though there is no one-to-one correspondence between the number of messages and the average success, their graphs do show some sort of mirroring on the micro level. The interrelationship of innovation force and average success is not well visible in Figure 4, because of the



**Fig. 5.** Simulation run of a  $3 \times 3^*$ -game with innovation in a 3-agents population. Comparison of communicative success and force of innovation; number of simulation steps at  $x$ -axis.

coarse scaling of the force of innovation value. Figure 5 shows the force of innovation and the communication success between step 50 and 350 of the simulation run, already depicted in Figure 4, whereas the force of innovation value is 20 times more fine-grained. Here the interrelationship between both values is clearly recognizable, one measure's peak is simultaneously the other measure's valley. Admittedly the mirroring is not perfect, but it improves by increasing the number of agents.

## 5 Conclusion and Outlook

Let's recap: We started out with comparing Roth-Erev and Bush-Mosteller reinforcement, finding that Bush-Mosteller yields better results for repeated signaling games. Extending Bush-Mosteller with lateral inhibition lead to even better results, but far from perfect. And results were even worse for multi-agent account with 3 or 5 agents: with increasing  $n$  less agents develop a signaling language in the first place, especially pooling strategies turned out to be a common outcome. In a next step we extended the classical Bush-Mosteller reinforcement by adding negative reinforcement and therefore achieving lateral inhibition, innovation and extinction. We found that these tweaks result in perfect communication between 3 agents in  $n \times n^*$ -games for  $n < 8$  and between 5 agents for  $n < 6$ , since higher values for  $n$  or the number of agents require much higher runtime that exceed our limit. Especially the force of innovation seems to be responsible for this achievement, since it makes sure that new messages are introduced when communication is not successful, while the combination of negative reinforcement and lateral inhibition takes care of all unused or useless messages to become extinct. Consequently, the result is an agreement on one single perfect signaling language with no other messages that might interfere.

The purpose of this direction of research is mostly about finding reasonable extensions of simple learning algorithms that lead to more explanatory results, assuming that more sophisticated learning dynamics might be more adequate to eventually describe human language behavior. We think the extensions we introduced in this article are of that kind, especially negative reinforcement, since we're rather certain that failure has a learning-effect, and innovation and extinction, because it seems unreasonable to

assume that all messages are available right from the start and that everything is kept in the lexicon, even if it has only once been successfully used. Further research in this direction should clarify how memory restrictions could be modeled and how sender and receiver roles of one agent should influence each other. What remains to be shown is that our results in fact hold for higher numbers of agents and states. It would further be interesting to see what influence different, again more realistic network-types (say small-world or scale-free networks) have on the results and what happens if two or more languages interact.

## References

- CE1. Clarke, F., Ekeland, I.: Nonlinear oscillations and boundary-value problems for Hamiltonian systems. *Arch. Rat. Mech. Anal.* **78** (1982) 315–333
- L1. Lewis, David: *Convention*. Cambridge: Harvard University Press (1969)
- B1. Barret, Jeffrey A.: The Evolution of Coding in Signaling Games. *Theory and Decision* **67** (2009), pp. 223–237
- BZ1. Barret, Jeffrey A., Zollman, Kevin J. S.: The Role of Forgetting in the Evolution and Learning of Language. *Journal of Experimental and Theoretical Artificial Intelligence* **21.4** (2009), pp. 293–309
- BM1. Bush, Robert, Mosteller, Frederick: *Stochastic Models of Learning*. New York: John Wiley & Sons (1955)
- HH1. Hofbauer, Josef, Huttegger, Simon M.: Feasibility of communication in binary signaling games. *Journal of Theoretical Biology* **254.4** (2008), pp. 843–849
- HSRZ1. Huttegger, Simon M., Skyrms, Brian, Rory, Smead, Zollman, Kevin J.: Evolutionary dynamics of Lewis signaling games: signaling systems vs. partial pooling. *Synthese* **172.1** (2010), pp. 177–191
- HZ1. Huttegger, Simon M., Zollman, Kevin J.: Signaling Games: Dynamics of Evolution and Learning. *Language, Games, and Evolution*. Ed. by Anton Benz et al. *LNAI 6207*. Springer (2011), pp. 160–176
- M1. Mühlenbernd, Roland: Learning with Neighbours. *Synthese* **183.S1** (2011), pp. 87–109
- MF1. Mühlenbernd, Roland, Franke, Michael: Signaling Conventions: Who Learns What Where and When in a Social Network. *Proceedings of EvoLang IX* (2011)
- RE1. Roth, Alvin, Erev, Ido: Learning in extensive-form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behaviour* **8** (1995), pp. 164–212
- S1. Skyrms, Brian: *Signals: Evolution, Learning & Information*. Oxford: Oxford University Press (2010)
- W1. Wagner, Elliott: Communication and Structured Correlation. *Erkenntnis* **71.3** (2009), pp. 377–393
- Z1. Zollman, Kevin J. S.: Talking to neighbors: The evolution of regional meaning. *Philosophy of Science* **72.1** (2005), pp. 69–85.