

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Modelling Human Communication as a Rejection Game

Permalink

<https://escholarship.org/uc/item/15g9k8c0>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Çapari, Artemis A., BSc
Sbardolini, Giorgio

Publication Date

2021

Peer reviewed

Modelling Human Communication as a Rejection Game

Artemis Çapari (artemiscapari@gmail.com)

Department of Computer Science, University of Amsterdam
1098 XH Amsterdam, The Netherlands

Giorgio Sbardolini (g.sbardolini@uva.nl)

ILLC and Department of Philosophy, University of Amsterdam
1090 GE Amsterdam, The Netherlands

Abstract

We present a computational model of a communicative interaction between a speaker and a hearer in a signalling game to study the emergence of compositionality. The partly cooperative/partly competitive interaction formalises essential aspects of human communication, with the goal of approximating human language use and analysing it by means of simulation studies. The basis for the interaction is a *Rejection Game* inspired by Robert Stalnaker’s account of conversation. This basis is then enriched to accommodate for compositional combinations of signals. The overarching goal is to study the emergence of stable meanings in a conversational model that allows for cooperative as well as conflicting players. Compositionality poses a challenge, as it allows players to exchange misleading signals that are only partially truthful. Furthermore, personal attributes integral to human decision making such as sympathy and trust were implemented in the model as adjustable parameters, thus providing the opportunity to create and study individuals with different ‘personalities’. Equilibria are obtained by a form of reinforcement learning with finite memory, and studied parametrically in terms of sympathy and trust. The model was able to substantiate some widely confirmed notions about human communication such as the correlation between truthfulness and trust in communication, demonstrating the possibility of correlation between aspects of linguistic cognition and social aspects of language use.

Keywords:

Compositionality of Language; Signalling Theory; Rejection Game; Multi-Agent Interaction; Reinforcement Learning;

Introduction

According to an influential tradition in cognitive science, language evolved under communicative pressure, balancing cognitive simplicity and expressive completeness (Skrms, 2010; Lewis, 1969; Goodman & Stuhlmüller, 2013). This paper provides a reinforcement learning model for the development of a basic form of combinatorial syntax, a trait common to all human languages. The communicative interaction is formalised as a *Rejection Game*, an abstract model of Robert Stalnaker’s theory of conversation (Stalnaker, 2014; Incurvati & Sbardolini, ms). A speaker S sends a message that may or may not be true, which the hearer H then may accept or reject. However, S prefers to avoid rejection while H prefers to accept true statements and reject false ones. These agents are faced with strategic choices when sending or accepting a message.

In this setting, compositionality poses a special challenge not reducible to combinatorial complexity. A complex message might be only partially true relative to the initial state.

For example, ‘Red Tree’ is partially true if the initial state is a green tree. Could a speaker lure the hearer into accepting a complex message which is only partially true? If so, compositionality would quickly disrupt the ability of interacting agents to develop a robust correlation of signals and meanings. Nevertheless, compositionality is essential to human communication, as it allows finite cognitive systems to express an infinity of meanings. A compositional language, as any human language, contains a fairly limited amount of words, but these words can be combined in different ways to have different meanings. Our model will thus include this fundamental aspect of language.

To illustrate the game as a real-life situation, let us consider the following example: Suppose S is a salesperson, and H is a customer. The customer wants to buy only good products, and the salesperson wants the customer to just buy products. S has to choose between telling H the truth about the products’ quality or not necessarily the truth but something that will compel H to buy. This shows how players are required to make strategic choices when their preferences do not align. The special challenge posed by compositionality in this scenario may then be understood as the challenge posed by misleading “package deals”: a sale of complex items some of whose parts may be good and some not. Such deals complicate the decision of the customer: to buy a package for its good component even if there is a bad component, or rather to walk away? Games of this kind are applied to the study of persuasion (Kamenica & Gentzkow, 2011; Pitchik & Schotter, 1987), but the abstract features of the Rejection Game represent a wide class of interactions in ordinary life.

Despite the potential for conflict, human players have a tendency for cooperation. This has been documented in several empirical studies and accounted for in theoretical approaches, often on the base of interactions akin to the Rejection Game (Cameron, 1999; Güth, Schmittberger, & Schwarze, 1982; Bicchieri, 2006). In the example above, the sender does not get rewarded less for failing to be truthful, as she might believe an untruthful message is more likely accepted. In contrast, a human speaker might sympathise with the hearer and be more inclined to tell the truth to mutual benefit. This does not come naturally to an artificial agent, as the best strategy might not be the most “benevolent”. We adopted a differential benefits model of cooperation by implementing an additional motive to be truthful, called *sympathy* (Sally, 2000, 2003).

Finally, a learning method that is the most similar to how humans would learn strategies must be developed. For the model to develop the expected behaviour, the learning method must require that the players do not base their strategies only on what happened in the previous round of the game. The players should keep in mind what the overall strategies of the opponents are, how the opponent responded to the player’s own strategies, and the ability to recognise that different situations require different strategies. We adopted a reinforcement learning model in which, realistically, players have a finite memory of previous interactions.

Our discussion will be focused on the following points:

- The evolution of language involves a complex web of forces sometimes pulling in different directions. Under which conditions do we observe the emergence of a compositional language in a conversational interaction?
- Compositionality is a fundamental aspect of human language, but it has the potential to disrupt the interaction between speaker and hearer. How does compositionality affect cooperative signalling behaviour?
- Multiple repetitions of the game allow for the agents to ‘learn from experience’. How can we implement plausible learning dynamics?

In the next section, a theoretical foundation for the present research is established. Methods and simulation design will then be discussed, followed by results and evaluation. Finally, we will discuss further directions for improving the model.

Theoretical Framework

The Rejection Game

The decision-theoretic structure of the Rejection Game is illustrated in Figure 1 (Incurvati & Sbardolini, ms). The game begins with a set of alternatives of which Nature selects one at random as the true state. The first player, a speaker S , has some evidence that indicates Nature’s choice. In Figure 1, the mark indicates that State B is the true state. S then chooses to send a message, A or B . S has an unconditional preference: her goal is to get her message accepted, whether she chooses to be truthful and send message B , or not, and send message A . The second player, the hearer H , does not have access to the initial evidence, but her task is to accept truths and reject falsehoods. In the simplest non-trivial setting, the players’ environment consists of a language $\mathbf{L} = \{A, B\}$ and $States = \{A, B\}$.

A Rejection Game can be considered a hybrid game, insofar as the speaker decides which game the hearer will play: cooperation or conflict (Incurvati & Sbardolini, ms). Table 1 represents a numerical payoff distribution matrix for the Rejection Game. The game becomes a coordination game if S sends a true message, as H wins if and only if S wins. However, if S sends a false message, the game becomes a zero-sum game: H wins if S loses and *vice versa*. Thus there are two salient equilibria among many: one in which S tells the truth

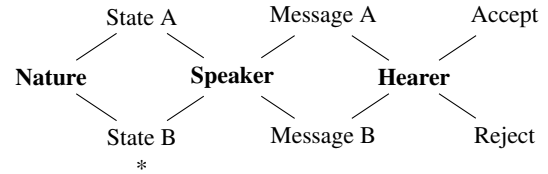


Figure 1: Decision tree of a 2x2 Rejection Game

in every state and H always accepts, and on in which S signals randomly and H always rejects.

Table 1: Hybrid Game: Payoff distribution

True State: A	Hearer		
	Accept	Reject	
Speaker	A	1,1	1,0
	B	0,0	0,1

In the basic model represented in Figure 1, the players are not yet capable of combining different symbols into meaningful wholes, such as ‘Red’ and ‘Tree’ into ‘Red Tree’. A complex signal AB is obtained by concatenation of simpler signals A and B . An advantage of compositionality is that it enables the artificial agent to form a larger number of expressions with a higher complexity, while requiring fewer different symbols. Thus, compositionality allows linguistic agents to satisfy a demand for expressive completeness within the computational limits of a finite mind. The simulation model will implement a rudimentary version of compositional combination of symbols into meaningful wholes. Additional complexity could be added as well in more sophisticated models of conversation. Some comments on this are included in the last section.

Method and Design

Simulation Model

In a simulation, the Rejection Game is played over multiple (at least 200) rounds. A state is selected at random each round to represent the truth: the true state \mathbf{T} . The value of \mathbf{T} is saved for each round to serve as ‘experience’ for future rounds. The program then chooses the payoff matrix \mathbf{P} corresponding to state \mathbf{T} , as the rewards are different for each state. For instance, if $\mathbf{T} = A$ and S chooses to send message ‘ B ’, H would have to reject S ’s message to receive a positive reward, while H should accept message ‘ B ’ when $\mathbf{T} = B$ in order to win. The hearer’s payoffs are thus based on \mathbf{T} . Payoffs for the speaker are also state-dependent, as explained below.

During each round r , two strategy distributions are calculated: one for the speaker S and one for the hearer H . The strategy distribution for the very first round is simply a random multinomial distribution. After the first round, strategy distributions are calculated by a learning function which takes the payoffs \mathbf{P} for state \mathbf{T} and the strategy distribution from previous round $r-1$ to calculate the strategy distribution for

the current round r . The model allows for different learning classes to be implemented, but for consistency across the present study we adopted a version of ‘Moran learning’: a form of finite memory reinforcement learning which we discuss below.

Compositionality

The function for creating a compositional language takes as inputs an initial list of symbols with length n , and a parameter for the maximum length m of the compositional symbols, and returns all possible permutations with repetitions up to and including that given length. The length s of the compositional language is of size $s = \sum_{x=1}^m n^x$.

With $[A, B]$ as the initial list of symbols and $m = 2$, the function will combine the set of symbols into a language that is comprised of the words $\mathbf{L} = \{A, B, AA, AB, BA, BB\}$. The program can take on any list of symbols and any length, but with the objective of illustrating how the model operates, any example below regarding the model itself will be assumed to have the language \mathbf{L} unless stated otherwise.

Rewards

Each player has their own payoff matrix, consisting of separate matrices for each true state which contain the payoffs within that state. S has as many strategies as there are words in the language, whereas H has twice as many strategies because she can choose to accept or reject each possible message. The dimensions of the matrices of each player are thus: true states \times S strategies \times H strategies (hence, $6 \times 6 \times 12$ with language \mathbf{L}). The payoffs are adjustable and different values are required for the parameters depending on which learning method is used. With language \mathbf{L} and Moran Learning, a payoff of 100 was found to be ideal for the speaker as well as the hearer and this payoff will be used in the tables and examples of this section.

Speaker Considering that the speaker’s goal is to get her message accepted and the rewards for acceptance are equal for both truthful and untruthful messages, S is easily inclined to be untruthful. Additionally, the first iteration is chosen from a random distribution that is weighted by the rewards, resulting in an even distribution. This entails that the probability that an untruthful message rather than a truthful one is chosen is larger: $(s - 1)/s \geq 1/s$. This behaviour however is not in accordance with how humans typically behave. Psychological evidence shows that people not only pursue their own material self-interest, but express altruistic behaviour as well, especially when caring about their opponent does not hurt them (Rabin, 1993; Cameron, 1999). We assume that a speaker would prefer to avoid being untruthful and hurting the hearer. The human tendency to sympathise with the opponent has a large influence on the players’ strategies and should thus be implemented in the simulation model. In incorporating a bias for fairness in social interactions, we follow a recent trend in many studies in the social sciences (Sally, 2000; Bicchieri, 2006). In practice, we implement sympathy as a dif-

ference on outcomes. S is better off from telling the truth and worse off for not doing so. The value of sympathy is denoted as parameter λ with $0 \leq \lambda \leq 1$. With $\lambda = 1$, S is extremely sympathetic. With a material reward of $a \in \mathbb{R}$, payoffs for S in case of acceptance are calculated as follows:

$$\text{Truthful Message Accepted} = a + a \cdot \lambda$$

$$\text{Untruthful Message Accepted} = a - a \cdot \lambda$$

The additional rewards/punishment proportional to λ need not be considered as material, but could be psychological, indicating how a human would feel better if they helped the other player or guilty for deceiving them.

Hearer Payoffs for H are based on the similarity between the message and the truth. Partial rewards are given for accepting or rejecting a partially true message. In order to assess similarity, two aspects need to be considered: which tokens occur and the order in which they occur. For the purpose of illustrating the importance of these two aspects, consider the sentence ‘dog bites boy’ (ABC). The sentence ‘boy bites dog’ (CBA) is less similar in meaning to ‘dog bites boy’ than for instance ‘dog bites cat’ (ABD) or simply ‘dog bites’ (AB). For the difference between ‘dog bites boy’ and ‘dog bites cat’, the similarity measure must be token based. As can be seen in Table 2, the Ratcliff-Obershelp distance (Ratcliff & Metzener, 1988) from the Python TextDistance library Module was found to provide the most appropriate balance between the importance of the meanings of the individual symbols in the sentence and the importance of the sequence in which they are ordered, compared to some prominent alternatives. This approach is promising but there are limits: the Ratcliff-Obershelp distance is a sequence-based distance measure and therefore does not fully account for the content of tokens.

Table 2: Similarity Measures

Truth	Message	Rat-Ober	Cosine	Levenshtein
ABC	ABD	~0.66	~0.66	~0.67
ABC	AB	0.8	~0.75	~0.67
ABC	CBA	~0.33	1	~0.31

Moreover, the hearer’s payoffs depend on a parameter that expresses the level of *trust* a hearer has for the speaker. The higher the trust the more fulfilling the interaction will be. As some individuals are naturally more trusting than others, this personal attribute should be represented in the model and is denoted by ϕ with $0 \leq \phi \leq 1$. With $\phi = 1$, the hearer is extremely trusting. To summarise, the hearer’s payoffs are calculated as follows: the sender’s message is M , \mathbf{T} is the true state, $a \in \mathbb{R}$ is the material reward for the action, and `normalised_ratcliff_ober` is the normalised Ratcliff Obershelp distance function from the Python TextDistance library.

$$\text{Similarity} = \text{normalised_ratcliff_ober}(M, \mathbf{T})$$

$$\text{Accepting } M = a \cdot \phi \cdot \text{Similarity}$$

$$\text{Rejecting } M = a \cdot \phi - (a \cdot \phi \cdot \text{Similarity})$$

Learning Method

Moran Learning A round of the game determines how the agents will play at the next round. We rely on a version of basic Roth-Erev reinforcement learning called ‘Moran learning’ (Roth & Erev, 1995; Moran, 1962). The Moran learning function implemented in our model requires a constant memory of accumulated rewards. The reward $V(n, x, X)$ at time n for playing x in state X increases if x was successful given X , or remains the same otherwise. Total rewards are constant, so that the individuals’ memory of past interactions doesn’t simply grow indefinitely. The probability that an agent chooses x in X at time n is given by the ratio between the rewards for x and the total rewards.

$$p(x|X)(n) = \frac{V(n, x, X)}{\sum_y V(n, y, X)}$$

Wins and losses are defined by the Rejection Game, in different ways for S and H as explained above. Rewards then accumulate for both players by two conditions.

$$V(n+1, x, X) = \begin{cases} V(n, x, X) + 1 & \text{if the agent won} \\ V(n, x, X) & \text{otherwise} \end{cases} \quad (\text{i})$$

$$V(n+1, x, X) = \begin{cases} V(n, x, X) - 1 & \text{if } V(n, t, s) > 1 \\ V(n, x, X) & \text{otherwise} \end{cases} \quad (\text{ii})$$

By (i), rewards accumulate if conducive to success; by (ii), the sum of rewards for all actions remains (roughly) constant.

Experience Function As a true state \mathbf{T} is picked at random during each round r , simply using the strategies from round $r - 1$ would not be useful as \mathbf{T} could have a different value and would thus require a different strategy. Therefore, an additional *Experience Function* was created. EF finds the last round in which \mathbf{T} was the truth and passes the probabilities of strategies from that round to the learning algorithm which uses it to determine the probabilities of strategies for the current round. Hence the agents decide which message to send based on \mathbf{T} . If a player does not yet have the appropriate experiences in a round r (i.e., in the first few rounds), EF simply uses the strategy distribution from round $r - 1$ until it does gain enough experience. This can be regarded as ‘trying out’ different strategies, similar to what humans might do when they are not yet sure of what the opponent’s strategies are. As they play, they learn from their experiences and are able to make more informed choices.

The Moran learning function takes the strategy distribution provided by EF as an argument, and calculates the ‘weight’ of a strategy, depending on expected payoffs. The algorithm iterates through a certain number of time steps and during each time step two rewards are chosen at random: one is replicated and one eliminated. The higher the weight of the strategy that the individual has chosen, the more likely it will be reinforced. These correspond to conditions (i) and (ii) respectively in the equations above.

Table 3: Summary of the previously defined variables.

Variable	Meaning
\mathbf{T}	True State
\mathbf{P}	Payoff Matrix
\mathbf{L}	Language, consists of (combinations of) words
S	Speaker/Sender
H	Hearer/Receiver
A	Word, combination of words is in language \mathbf{L} and can result in a state \mathbf{T}
B	Word, combination of words is in language \mathbf{L} and can result in a state \mathbf{T}
r	A round in the rejection game
a	Reward
λ	Speaker’s level of sympathy
ϕ	Hearer’s level of trust

Results and Evaluation

We begin by considering the effects of the Experience Function. For this we set \mathbf{L} as language, sympathy $\lambda = 1$, trust $\phi = 1$, memory size of 200 and $200 \cdot 6 = 1200$ rounds of the game. Figures 2(a) and 2(c) show the proportion of strategies for the Speaker for each state along with the linear regression of truthfulness across all states. Figures 2(b) and 2(d) show the Hearer’s average response of all states combined along with the standard deviation.

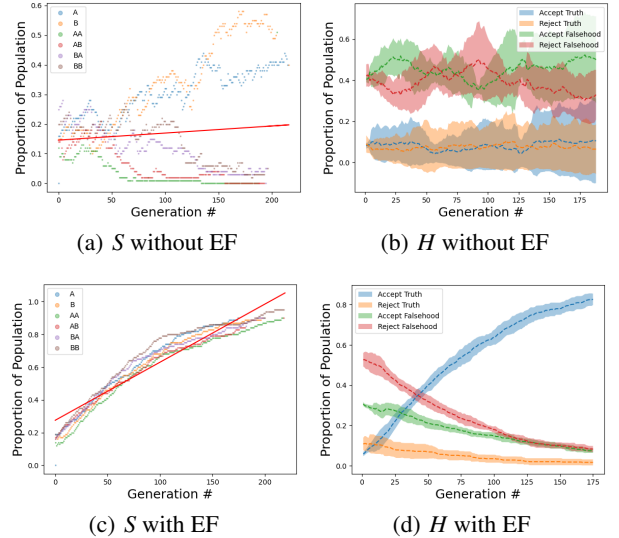


Figure 2: Influence of the Experience Function

The Experience Function improves the performance of the players (Figures 2(c) and 2(d)). The speaker learns to be truthful in each state and the hearer learns to accept the true message. Without EF, Figures 2(a) and 2(b) indicate that truthfulness and trust do not emerge robustly. Since in the model without EF the players based their strategies only on the previous round, and the true state \mathbf{T} is picked at random at each round, the players are quite literally playing randomly. Consequently, failure of optimization is to be expected.

Truthfulness and Trust

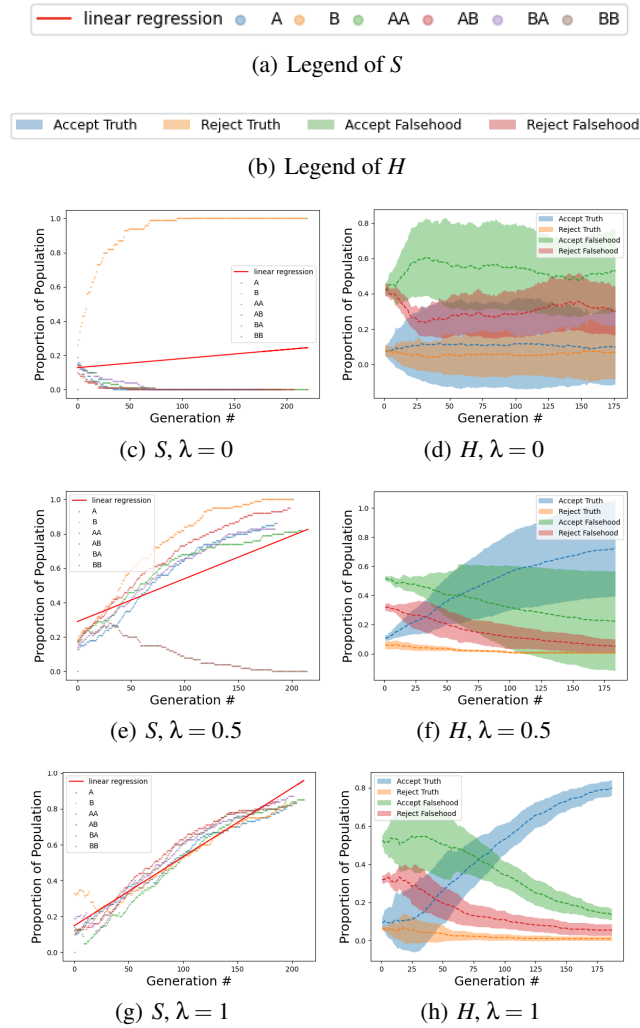


Figure 3: Performance Based on different λ -values

Next, we assess the impact of sympathy. Figure 3 shows the results of simulations across various values of λ . As λ increases, the speaker learns to become more truthful and the hearer in turn learns to accept the speaker's messages. With no sympathy, the computer agent simply relies on chance, as can be seen in Figure 3(c). Here there are no rewards for truthfulness and no punishments for deception. This entails that it is far more likely that an untruthful strategy is chosen ($\pm 5/6$) rather than the truthful strategy ($\pm 1/6$). Consequently, in only one state out of six (state *B*) the speaker could be said to learn efficient behaviour (see Figure 3(c)), but this seems more a matter of luck.

Trust also matters. In Figure 4 we have $\lambda = 1$, yet unlike previous examples, the hearer is set to not receive any partial rewards: $\phi = 0$. She only receives a reward when she accepts a message that is equal to the truth and rejects anything other than the exact truth, making her more inclined to rejection. This approximates the behaviour of an individual who is nat-

urally mistrusting.

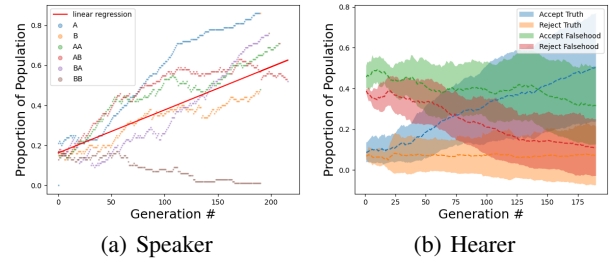


Figure 4: Influence of Trust with $\lambda = 1$ and $\phi = 0$

It becomes apparent that the speaker's sympathy on its own does not ensure efficient communication. Though we have a fully sympathetic speaker with $\lambda = 1$, they do not learn to be completely truthful at the end of the simulation when the hearers do not trust them (i.e. $\phi = 0$). Here, the hearer is more inclined to reject and the speaker's learning trajectory is slower. This could be explained by the fact that the chance of rejection from the hearer is higher when ϕ is lower. Hence, these results imply a direct correlation between truthfulness and trust in the development of language.

Compositionality

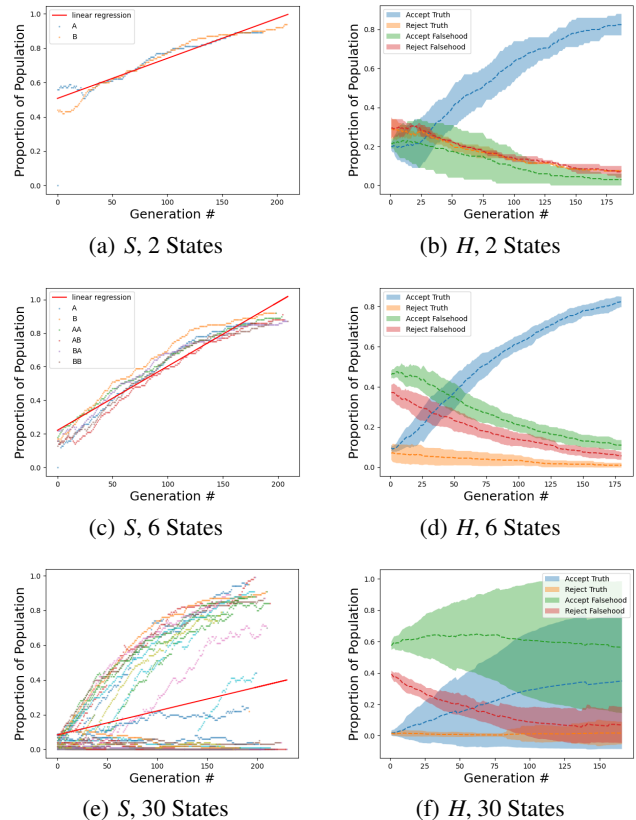


Figure 5: Performance Based on Language Size

The graphs in Figure 5 demonstrate that as the number of strategies to choose from rises, it becomes more difficult to learn the correct strategies. Each game in Figure 5 is played for $200 \cdot (\text{number of states})$ rounds, with $\lambda = 1$ and $\phi = 1$. In Figures 5(a) and 5(b), with two states, efficient communication is obtained at the fastest rate. As the number of states increases, it becomes more difficult to learn the appropriate strategies. In Figures 5(c) and 5(d), the optimal strategies are learned, but at a slower rate, whereas in Figures 5(e) and 5(f), with 30 states, they are sometimes not learned.¹ Due to the fact that there are more states, there are more partial rewards to be given to *H* and thus she becomes more likely to accept messages in general. As Figure 5(f) shows, *H* accepts almost any message. As a result, *S* does not learn truthfulness and *H* accepts falsehoods.

Discussion

From the obtained results, it can be concluded that the experience function has allowed the agents to ‘learn from experience’ over the course of multiple repetitions of the game. In combination with Moran learning, a reinforcement learning algorithm with finite memory (Moran, 1962), it has shown to be a plausible learning dynamic. However, it is to be noted that the experience function could be improved, as agent *S* only takes her own optimal strategies for each state **T** into account, without considering the response of *H*. As the truth is not necessarily the message that will most likely be accepted by *H*, *S* must always consider the hearer’s response.

While rejection should have the purpose of update-blocking and thus discourage the speaker from repeating their strategy, the way in which the experience function is implemented slightly negates the power of rejection that the hearer has. Although the speaker’s behaviour is favourable as a result, the speaker is indirectly forced into being truthful. It can thus not yet be considered to be an optimal learning method for approximating human communication.

Another finding is that the behaviour of the agents in this model is in accordance with an aspect of human communication that has been widely recognised, namely that the speaker’s truthfulness and the hearer’s trust are correlated (Lewis, 1969). By assigning different personality traits to the interlocutors and analysing their influences on communication, a correlation was found between sympathy from *S* and trust from *H*. Truthful communication was not obtained with a sympathetic speaker and a mistrusting hearer or a non-sympathetic speaker and a trusting hearer. It can thus be stated that efficient communication emerges only when there is sympathy as well as trust.

Furthermore, as the complexity of the language increases, *H* becomes more likely to accept messages of *S* due to the fact that *S* can now send partial truths. The ability for a speaker to send partial truths that are accepted by the hearer demotivates her from telling the exact truth, and thus optimal strate-

gies are seldom attained. As the number of possible strategies increases, more noise is created and memorising the appropriate strategies becomes more difficult. So we find that compositionality does have a disruptive force on cooperative behaviour, but that its negative impact can be alleviated for some specific parameter setting.

The implementation of compositionality in the model has allowed the model to show behaviour approximating that of human beings. It provided the interlocutors with better means of communication along with the complications that a more complex language brings. Compositionality in general is a much more complex phenomenon (Piantadosi, Tenenbaum, & Goodman, 2016). In this model compositionality in human language has been formalised as follows: the compositional language consists of the permutations with repetition of a finite number of symbols. These are then compared by similarity of sequence and content to signify their meanings. Although rudimentary in its resemblance of natural language, the model does accommodate for complex signals.

These findings are consistent with well established theories about the interactive origins of our linguistic abilities, and the role of communication in the development of language. The findings are also illustrative of some of the difficulties that may arise when combining several aspects of language evolution in a single model, specifically the potential conflict between cooperative speaker/listener behaviour (truthfulness and trust) and a compositional language. We find that the main value of this research lies precisely in its ambition to integrate several threads in research on language evolution within a unique model of communication, in which discussion of simple facts concerning efficient communication may tell us something about language that we do not yet know or have not yet proven.

Conclusion

Human communication was approximated in the simulation model by implementing a language that allowed for communication with complex signals through compositional signalling. With a language such as this, parameter setting shows that the agents expressed behaviour similar to that of human agents. As *S* is equipped with the tools for more complex communication, she was able to get her message accepted by *H*, even when telling only partial truths. Sometimes, this resulted in the agents not being able to learn their optimal strategies. Furthermore, a learning dynamics such as Moran learning proves to be a mostly successful method for mimicking the way in which human interlocutors learn to communicate when provided with the appropriate experiences as their learning input. Finally, the findings demonstrate that the model has the ability to tell us about aspects of human communication and how personal attributes such as an individuals trusting nature or their ability to sympathise with others affects the manner in which they communicate. While the correlation between truthfulness and trust is well established, it was not intentionally implemented in the model, showing

¹The legend of *S* is not shown as there are too many, but words A and B were combined with a max length of 4 with 40 possible states.

that the model can confirm knowledge about human communication and potentially even provide us with new knowledge.

Although the model has many applications, it would need several improvements before being able to realistically depict human interaction. A current limitation is that the agents learn about each other “by brute force”, being exposed to trials and errors over repeated interactions. As the speaker currently bases their strategy on which state is true without considering the hearer’s response, the expectation function from the learning method could be improved as well. This could be achieved by further specifying a reasoning method in which the speaker takes the hearer’s response into account as well. In this respect, the agents in our model do not represent each other, and do not make choices based on a probabilistic assessment of the opponent’s knowledge and behaviour. More sophisticated models of cognition can do this, such as Rational Speech Acts models of linguistic interactions (Goodman & Stuhlmüller, 2013; Goodman & Frank, 2016). One possible improvement of our study could be to integrate this research in a more articulate Bayesian framework.

Another improvement could concern the language used in the model. The model takes into account both sequence similarities and token similarities when comparing sentences. However, weights should be added to the signals to signify the importance of each word, as some words contribute more to the meaning of a sentence. Weights on similarity of words could also be used to recover lexical approximations. For example, ‘cat bites dog’ is not exactly true in a “cat scratches dog” state, but it is also not as far from truth as ‘cat befriends dog’. These subtleties of meaning ought to be accounted for in a proper discussion of efficient communication for a compositional language, but they also appear to call for a wider methodology than we could employ in our study.

In order to truly evaluate whether the behaviour of the agents resembles that of humans, future work could include an empirical evaluation of the model with human subjects. In this evaluation, human subjects would play the rejection game in the same manner as in our simulation model. However, as humans are not able to play 1200 rounds in succession, the evaluation would naturally consist of fewer repetitions. This already highlights a drawback in the simulation model as it takes many repetitions before the players learn their strategies while human subjects are not even able to continue playing for that many sessions. Therefore, the speed at which the agents learn could also be improved.

Although the model should be evaluated with human data, it cannot be denied that studying communication between agents in a computer simulation is far more convenient than studying human subjects. If further improved upon, a computer program such as this could serve numerous different purposes considering it can play thousands of rounds in succession whilst efficiently saving the data from each round. As the model is implemented with the ability to amplify certain personality traits, one possible study could for instance be communication at a more individual level as well.

Acknowledgments

This work has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement No 758540) within the project *EXPRESS: From the Expression of Disagreement to New Foundations for Expressivist Semantics*.

References

- Bicchieri, C. (2006). *The grammar of society*. Cambridge: Cambridge University Press.
- Cameron, L. A. (1999). Raising the stakes in the Ultimatum game: Experimental evidence from Indonesia. *Economic Inquiry*, 37(1), 47-59.
- Goodman, N., & Frank, M. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20, 818-829.
- Goodman, N., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, 5, 173-184.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization*, 3, 367-388.
- Incurvati, L., & Sbardolini, G. (ms). *The rejection game*.
- Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101, 2590-2615.
- Lewis, D. (1969). *Convention*. Cambridge, MA.
- Moran, P. (1962). *The Statistical Processes of Evolutionary Theory*. Oxford: Clarendon Press.
- Piantadosi, S. T., Tenenbaum, J. B., & Goodman, N. D. (2016). The logical primitives of thought: Empirical foundations for compositional cognitive models. *Psychological Review*, 123, 392-424.
- Pitchik, C., & Schotter, A. (1987). Honesty in a model of strategic information transmission. *American Economic Review*, 77, 1032-1036.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *The American economic review*, 1281-1302.
- Ratcliff, J. W., & Metzener, D. E. (1988). Pattern-matching-the gestalt approach. *Dr Dobbs Journal*, 13(7), 46.
- Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8, 164 - 212.
- Sally, D. (2000). A General Theory of Sympathy, Mind-Reading, and Social Interaction, with an Application to the Prisoners’ Dilemma. *Social Science Information*, 39, 567-634.
- Sally, D. (2003). Risky speech: behavioral game theory and pragmatics. *Journal of Pragmatics*, 35(8), 1223-1245.
- Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.
- Stalnaker, R. (2014). *Context*. Oxford University Press.