Emergence of Communication in Coordination Games with Signaling Strategies

Zahra Moradi and Anil Yaman

Vrije University Amsterdam

Abstract. Communication is an important factor in the evolution of human beings. Individuals are constantly interacting to achieve both their self-fulfilling and shared goals on a daily basis. To this end, a uniform language has emerged among humankind in order to better express their desires and communicate together to convey their thoughts. Similarly, intelligent agents that exist in a common environment could benefit from this form of communication. This paper analyzes the emergence of language in multi-agent systems in a game environment.

Keywords: Communication · Language · Intelligent systems · Signaling Game· Coordination Game · Epsilon-Greedy Algorithm · Language Game.

1 Introduction

The emergence of language has been a topic of fascination for linguists and anthropologists for centuries. The question of how humans developed the ability to communicate complex ideas and emotions through language has led to numerous theories and debates. Today, as we witness the rise of artificial intelligence and machine learning, the study of language emergence has taken on new significance. With the development of sophisticated language processing algorithms, researchers are beginning to explore the similarities and differences between human and AI communication in various fields, including language game theory, evolutionary linguistics, and cognitive science. Spike et al. (2017) discusses several related works that investigate functional learned communication following the development of other human signaling conventions [1].

In this paper, we will explore the emergence of language and its connection to AI communication. Specifically, we will investigate the ways in which language has evolved over time and how this evolution has influenced our understanding of AI language processing. We will also examine the various theories of language emergence and how these theories inform our understanding of the development of AI language. We aim to answer the research question, "To what extent do agents agree about the semantics of the emergent language in coordination games, and how can this agreement be improved?" By exploring these connections, we hope to gain a deeper understanding of the ways in which humans and machines communicate and how this communication can be optimized for better outcomes. As discussed by Steels (2012), communicative success is a determining factor in the language selection and evolution [2]. Therefore, designing scenarios in which signaling affects the individual rewards of the agents is an important part of our research. In addition, we will explore different sub-research questions such as "Do the number of available agents affect the agreement in the emergent language?" and "What factors contribute to reducing the time required to achieve higher agreement?". This paper aims to study the emergence of language in a multi-agent game environment that requires proper coordination among the agents to be solved. To this end, a game scenario is designed with several available agents in which communication is the key element to achieve the best results.

2 Related work

There have been several research studies conducted to study emergent communication in artificial intelligent systems. The ability to transfer knowledge in a comprehensive scenario assists agents in achieving a common goal. Language plays a significant role in this communication, allowing us to analyze the similarities between human and intelligent systems. One influential theory that attempts to explain the emergence of communication is signaling game theory. Signaling game theory posits that communication between individuals evolves as a result of their shared interest in a particular goal. In this scenario, one individual serves as the sender of a message, while the other acts as the receiver. The sender aims to communicate information about the goal to the receiver, who then acts accordingly.

Language game theory posits that language emerges as a result of people engaging in repeated interactions that involve both communication and coordination with one another, as discussed by Nowak et al. (1999) [3]. Through these interactions, individuals develop a shared understanding of the meaning of words and phrases, leading to the development of a shared language system. Recent research in the field of language emergence has focused on using computational models to simulate the evolution of language. These models employ algorithms to simulate the process of communication and the evolution of language over time. One example of this type of research is the study of the iterated learning model conducted by various sources such as Kirby and Hurford (2002). It involves passing on a language from one generation to the next, allowing researchers to observe how the language changes over time [4].

Other studies have focused on the role of cultural transmission in the emergence of language. For example, a recent study by Tamariz et al. (2023) found that the transmission of cultural information plays a crucial role in the evolution of language [5]. The study used a computer model to simulate the evolution of language over time and found that the transmission of cultural information was essential for the emergence of a shared language system. Overall, the study of language emergence and its ties to signaling game theory and language game theory has led to a greater understanding of how communication systems evolve over time. Through the use of computational models and other research methods, scholars such as Kauhanen (2020) and Barrett and VanDrunen (2022) are gaining new insights into the mechanisms that drive the emergence of language and how it can be optimized for improved outcomes in AI communication [6] [7].

The communication scenario is inspired by the "Lewis signaling game" proposed by Izquierdo et al. (2019) [8]. The Lewis signaling game consists of two players with the roles of sender and receiver. The game environment can have multiple states corresponding to different situations, and the sender is aware of the current state of the game. Additionally, the sender communicates with the receiver by choosing from a fixed set of signals. The receiver is unaware of the game state and can only observe the signal. Each state of the game has a unique correct action that is of interest to both the receiver and the sender. The receiver needs to determine this correct action by analyzing the sent signal. The results of Lewis' proposed scenario show that even the simplest form of this signaling game has various Nash equilibria, leading to a situation where the best payoff for any given player is to continue with their current strategy. As a result, a global stasis is reached, and the game play of the agents does not change further.

Godfrey (2011) discusses the Lewis signaling game approach in more depth. In the Lewis signaling game, there are many equilibria, and Godfrey (2011) analyzes both an evolved signaling system that explores a variety of signals and the basic signaling game consisting of two signals [9]. The findings suggest that adaptive dynamics play an important role in the success rate of agents' communication. Additionally, starting with unsophisticated learning dynamics, we can observe the emergence of meaningful signals and successful communication using the basic system of the signaling game. This holds true for evolved systems with many signals as well, further demonstrating that meaningful communication can emerge without pre-existing meaning

Moreover, Franke (2008) investigates the semantic meaning in exchanged signals among agents in a non-cooperative conversation [10]. Their research takes an epistemic approach by using a sequence of iterated best responses (IBR) model and analyzing the focal meaning of the sent message. The IBR model relies on a sequence of best responses, starting with a purely semantic truth-only sender strategy. If the sender does not have an intentional incentive to deceive the receiver, the received message is also believed to be truthful. Furthermore, agents have limited resources for reasoning, which limits the depth of the agent's meaning examination iteration.

Lipowska and Lipowski (2018) conducted similar research on the emergence of communication in a multi-agent environment that is trained with reinforcement learning techniques [11]. The experiment involves several agents and objects, and the agents need to agree upon a referential name for each object through communication. Their model is inspired by the Pólya urn model, in which a ball is randomly drawn from a set that contains white and black balls. After the draw, the ball is returned to the set with an additional ball of the same color, infinitely. In Lipowska and Lipowski's model, the agents are similar to the ball example, meaning that in each iteration, agents will communicate in pairs to agree on a given name for an object. After the agreement, the weighted selection of the word is increased for the given object. They conduct experiments on two different settings: the single-object version, which consists of a single object and N agents, and the multi-object version, which consists of M objects and N agents. These settings are analyzed to observe the emergence of language among the population of agents.

In our method, all agents share a common interest in ensuring the maximum payoff over time. Each agent chooses a role and receives a payoff based on the current games. To achieve higher rewards, the agents need to communicate with each other and collectively decide on the desired matching selection pattern for the games. We hypothesize that learning the game as well as the communication itself will have a great impact on the agreement among the agents. Furthermore, different coordination games, such as symmetrical and non-symmetrical reward distributions, together with the size of the population, are expected to affect the agreement percentage of the group. This paper focuses on analyzing the emergence of language in the proposed multi-agent game scenarios and discusses the differences in semantics for each individual agent and how they can be unified among the group for maximum agreement.

3 Methods

The scenarios, in general, consist of a simple game in which participants choose a role and receive a reward. To ensure that the game's success depends on communication, the chosen roles are affected by the received signals, and the optimal reward is only achievable if all agents follow the game rules. At the beginning of each round, the sender and receiver are determined. The sender then chooses a role based on their internal preference (at the start of the game, all roles have an equal chance of being chosen until further adjustments are made). In each round of the game, the sender can communicate with other receivers by choosing to send a signal "W" if they have a signal that matches their current state (role). If the signal base is empty, they can choose to generate a new signal for their current state based on a probability value of "P-signaling".

It is important to mention that if this probability is not met, they will not generate a signal, and as a result, they will not send a signal. The receivers can then decide their action based on the information received. This phase is referred to as the communication phase. The reason for this design is to investigate the emergent language between the agents and analyze different semantics and interpretations. After the communication phase, all agents receive payoffs based on the games. Consequently, they will update the confidence level of the communicated signal based on the outcome of the payoffs for future rounds. It is important to note that solving the games relies on communication, meaning that without any signals, the receiver would always pick roles randomly, regardless of their value or the state of the game. This is to ensure that signaling correlates with the performance of the agents. We explore four different scenarios that slightly differ in their communication procedures, and each scenario is explained in more detail in the following sections.

3.1 Scenario 1 - Fixed Sender And Receiver(s) Without Value Based Role Selection

The games determine the payoffs of the agents based on their chosen roles. We construct Scenario 1 by taking inspiration from the Lewis signaling game [8]. In Scenario 1, there is one fixed sender throughout all rounds of the game, with the other agent(s) acting as the receiver(s). This scenario is divided into two different settings: one with agents learning the value of each role (which is determined by the games) and one without the learning process of the values of the roles.

In Scenario 1, each signal consists of three different segments. For both the sender and the receiver, the first segment is the symbol of the word, which agents use for communication and can be recognized by receivers if it exists in their signals. The second segment is interpreted differently depending on the agent: for the sender, this segment refers to their internal state in the game, which is the role they have chosen for the current round of the game; for the receiver, this segment refers to the action they must take upon receiving the symbol of this word, which is the role they need to choose for the current round of the game. If the receiver cannot recognize the symbol of the word, they will generate a new signal with the received symbol and assign a random action to the word. The last segment refers to the confidence level of the word for both the sender and the receiver, which determines the usefulness of the agent's signal. The structure of the signal is displayed in Table 1. The sender is the first agent to choose a role. This decision can be random, or the sender can alternatively learn the value of the roles using a Q-learning algorithm. We explore both settings to observe the effect of learning on the agent's game-solving capabilities.

Table 1: This table displays the structure of the signals and their interpretations for a sender and a receiver respectively for the scenario 1.

(a) Sender

Symbol	State of Sender	Confidence
W1	self.role	conf_lvl

(b) Receiver

Symbol	Action of Receiver	Confidence
W1	random_role	conf_lvl

In Scenario 1, the sender is randomly selected among the agents in the first round of the game and remains fixed throughout all subsequent rounds until the game finishes. Each agent starts with an empty initial signal base. The game begins with the sender randomly choosing a role for the current round, followed by a communication phase. During this phase, the sender can generate a signal for their current state with an initial confidence level of 0.5 and a probability of p-signaling = 0.1. The receiver(s) checks their own signal base for a matching symbol of the received signal. Since the signal base of all agents is empty at the beginning of the game, the receiver(s) add the received symbol to their signal set, assign a random action to it, and set an initial confidence level of 0.5. In later stages of the game, if the receiver(s) have a signal in their signal set that matches the symbol of the received signal, they will choose the role that is assigned to that symbol.

After the communication phase ends, all agents receive payoffs based on the games. The sender and receiver(s) then update the confidence level (c_level) of the communicated signal using a Q-learning algorithm. This algorithm relies on the agent's payoff for the current game round (agent_p) with a learning rate of lr_cf = 0.1, as shown in formula 1. Additionally, the confidence level of all signals for an agent decays at a rate of decay = 0.99 after each round of the game, as shown in formula 2. Furthermore, every 2 rounds of the game, all agents remove signals with a confidence level below 0.1. This ensures that old signals with no utility are discarded, making room for the generation of new signals.

$$c\_level = lr_{cf} * (agent\_p - c\_level) \tag{1}$$

$$c\_level = decay * c\_level \tag{2}$$

3.2 Scenario 2 - Fixed Sender And Receiver(s) With Value Based Role Selection

The main difference between value-based agents in Scenario 1 and non-value-based agents in Scenario 2 is that in Scenario 1, the sender chooses their role based on the learned value of the roles (role_v) throughout the game. To be more precise, each agent updates their internal value for the chosen role using a Q-learning algorithm that relies on their payoff for the current round (agent_p) with a learning rate of lr_v = 0.5 (see formula 3). Additionally, to ensure that agents explore all possible roles, the sender has an internal exploration probability of p_exploration = 0.1, allowing it to randomly select roles. This helps prevent exploitation in cases where the game has more than one optimal option.

$$role\_v = lr\_v * (agent\_p - role\_v) \tag{3}$$

We expect this scenario to be effective in solving simple game settings, such as situations where all agents should pick similar roles. However, in scenario 1, the communication is one-sided. This means that in a group of agents with a population of 3 or more, each receiver agent only communicates with the sender, and there is no direct communication between the receivers in the group. To overcome this limitation, we propose scenario 2, in which the sender and receiver(s) are randomly chosen every round of the game. This ensures that each agent has a chance to become the sender and prevents one-sided communication.

3.3 Scenario 3 - Shuffled Sender And Receiver(s)

Scenario 3 involves a random sender in each round of the game, with the other agent(s) acting as receiver(s). In this scenario, we introduce a new segment in the signal structure, resulting in each signal consisting of four different segments. These segments are interpreted in the same way for all agents.

The signal in Scenario 2 consists of four segments. The first segment represents the symbol of the word, which agents use for communication and can be recognized upon receipt. The second segment represents the agent's internal state in the game, corresponding to the role they have chosen for the current round. The third segment indicates the action the agent must take when they receive the corresponding symbol, representing the role they need to choose for the current round. If the agent cannot recognize the symbol, they will generate a new signal with the received symbol and assign both a random action and a random state to the word. The last segment denotes the confidence level of the word, determining its usefulness for the agent. The signal structure is displayed in Table 2. In this scenario, we consider only the setting where agents learn the value of each role, ensuring that they do not exploit local minimum solutions.

Table 2: This table displays the structure of the signals and their interpretations for a sender and a receiver respectively for the scenario 3.

(a) Sender

Symbol	State of Sender	Action of Receiver	Confidence
W1	self.role	random_guess	conf_lvl

(b) Receiver

Symbol	State of Sender	Action of Receiver	Confidence
W1	random_guess	random_role	conf_lvl

In the modified version of this game setting, each round follows a similar process to Scenario 1, with the exception that the sender is randomly chosen among the agents for every round of the game. After the communication phase concludes, all agents receive payoffs based on the games. The sender and receiver(s) then update the confidence level (c_level) of the communicated signal using a Q-learning algorithm that relies on their payoff from the current game round (agent_p) with a learning rate of lr_cf = 0.5, as shown in formula 1.

This change is primarily motivated by time limitations and computational constraints, aiming to reduce the number of game rounds required to solve the game. Introducing two random variables in the signal's structure necessitates this adjustment. The decay hyperparameter (decay) and the internal learning rate of agents for the values of the roles (lr_v) remain unchanged, as displayed in formulas 2 and 3. Furthermore, all agents will remove signals that have a confidence level below 0.25 after every 2 rounds of the game. This adjustment is again made to accommodate time limitations and computational constraints.

3.4 Scenario 4 - Pairs Of Sender And Receiver

This scenario aims to enhance agreement among agents. To achieve this, we randomly select a pair of agents in each round of the game and assign them as the sender and receiver. These two agents are the sole participants for that round, taking actions and updating their signal base accordingly. The signal structure used in this scenario is similar to that of scenario 3, as displayed in Table 2. The game settings and formulas employed for updating the signal base and learning algorithms remain identical to those of scenario 3.

To further analyze the overall agreement within the group, we consider a setup with only two available roles: A and B. Additionally, since the receiver makes random guesses for two segments of the signal base, we have $2^2 = 4$ possible signal creations when the signal base is empty, as displayed in Table 3. This means that there is a 25 percent chance of agreement between any two agents, or in other words, a perfect match for the signals. It is important to note that this percentage applies only to the pair of agents, not to the group as a whole. However, we expect the group agreement to be higher than this percentage when using scenario 4.

Table 3: This table displays the possible signals that the receiver can create with only two available roles.

(a) Receiver

Symbol	State of Sender	Action of Receiver	Confidence
W1	A or B	A or B	conf_lvl

3.5 Scenario 5 - Pairs Of Sender And Receiver Where The Sender Adjusts Its Signal Based On the True Action Of The Receiver

In addition, we analyze a case where the sender can observe the receiver's action and adjust its sent signal if the perceived action matches its guessed semantics. This scenario is identical to scenario 3, except for the addition of the sender's ability to observe the receiver's action. To explain in more detail, at the end of each round, the sender checks whether the true action of the receiver matches its signal segment labeled 'Action of Receiver' in the sent signal. If there is a successful match, the sender increases the confidence level of this signal by an additional 0.1. On the other hand, if there is no match, the sender decreases the confidence level of this signal by 0.1. We hope that this small modification in the enhanced scenario 1 will result in a faster agreement between agents.

3.6 Game Configuration And Hyperparameter Settings

We run 30 games, with each game consisting of 10,000 rounds. The payoff scheme is displayed in Table 4. These schemes refer to different games and how the agents will be rewarded based on their role selection.

We start with two available roles, A and B, and two agents. The rows and columns represent the sender and receiver, and the number in each cell represents the reward value for the agents if the corresponding roles are chosen for each round of the game. For each game, agreement is considered to match if the agents have identical signals for similar symbols with a confidence level higher than 0.4. The game is considered to have successful agreement if this matching is true for all agents in the group. To calculate the percentage of agreement over all 30 games, we simply divide the number of games that had successful agreement by the total number of 30 games.

Table 4: This table displays different games that the agents need to follow in different games to achieve the corresponding rewards.

(a) Game 1

	Receiver	
Sender	A	B
A	(5,5)	(0,0)
B	(0,0)	(5,5)

(b) Game 2

	Receiver	
Sender	A	B
A	(0,0)	(5,5)
B	(5,5)	(0,0)

(c) Game 3

	Receiver	
Sender	A	B
A	(0,0)	(0,0)
B	(5,5)	(0,0)

4 Results

In this section, we analyze the outcome of three runs with the games explained in Table 4. For each game, we display the achieved percentage of agreement and the agents' cumulative reward averaged over all 30 games.

Scenario 1 demonstrates an average agreement of 84 percent across all games. This is expected since the communication is one-sided, and the receiver only needs to match one random variable, namely their action, to the sender's signal. In "Game 1" and "Game 2," the agents quickly converge to the solution, and the average payoffs remain at 5 consistently. Consequently, the agents are able to solve "Game 1" and "Game 2" without difficulty. However, they struggle to solve "Game 3" due to the fact that the sender randomly selects a role at the beginning of each round. Nevertheless, the sender and receiver's signals indicate that the agents do agree on the correct role with matching semantics, but fail to solve the game due to the random role selection by the sender. We anticipate this issue to be resolved with value-based role selection. All results for Scenario 1 are displayed in Figure 1 and Table 5.

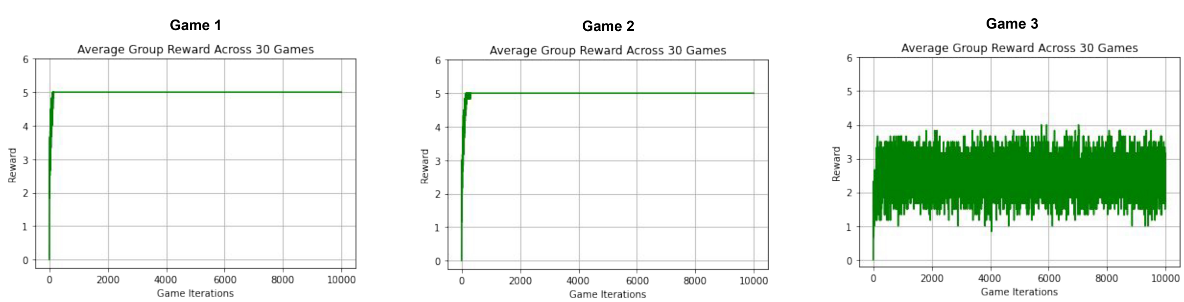

Fig. 1: This figure displays the results of different games for scenario 1.

Scenario 2 demonstrates an average agreement of 94 percent across almost all games. This is still expected since the communication remains one-sided, similar to Scenario 1. The agents are now able to solve all three games successfully. This is attributed to the agents learning the value of each role presented by the games, and the sender selecting the role with the higher value to achieve better payoffs. In "Game 1," the agents quickly converge to the solution, and the average payoffs remain consistently at 5. In "Game 2" and "Game 3," the agents also converge to the solution, but the average payoffs fluctuate slightly due to the added exploration rate, meaning that the sender occasionally selects the role with the lower value. All results for Scenario 2 are displayed in Figure 2 and Table 5.

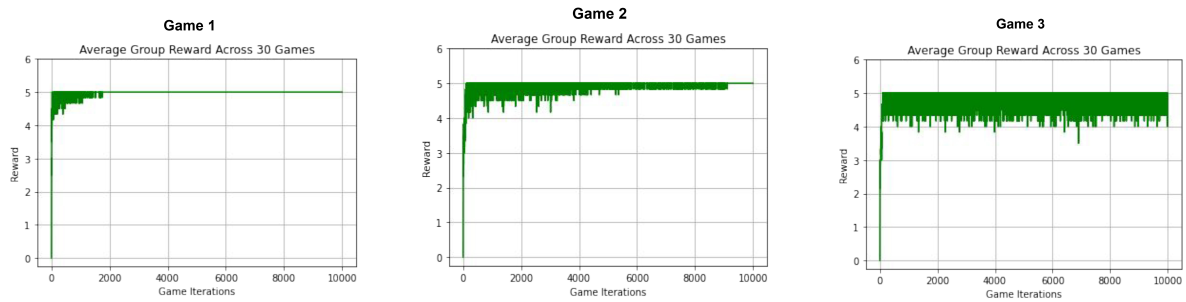

Fig. 2: This figure displays the results of different games for scenario 2.

Scenario 3 exhibits an average agreement of nearly 91 percent across all games. In all three games, the agents converge to the solution, but the average payoffs fluctuate slightly due to the added exploration rate. This means that the agents occasionally select the role with the lower value if they are chosen as the sender. It is important to note that the communication is no longer one-sided, and each member of the population has the potential to send signals. This is a promising result, but further iterations are required to achieve a higher agreement percentage, particularly as the number of agents in the group increases. This is because all agents need to coordinate their signal adjustments, which can be more challenging and computationally costly in a larger population. All results for Scenario 3 are presented in Figure 3 and Table 5.

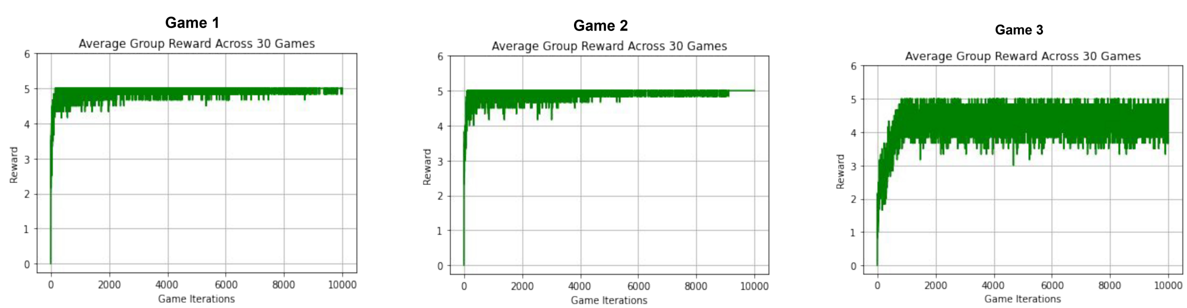

Fig. 3: This figure displays the results of different games for scenario 3.

Scenario 4 exhibits a similar average agreement of nearly 92.2 percent across all games. The plots display comparable results to Scenario 3. This similarity arises because, in a two-agent game, Scenarios 3 and 4 operate in a similar manner. The distinction lies in how they handle an increase in population size. In Scenario 4, two random individuals are selected as the sender and receiver, whereas in Scenario 3, one individual acts as the sender while the remaining agents serve as receivers. It is important to note that the averaged cumulative rewards pertain to the reward values of the pairs in each round and not to specific agents. All results for Scenario 4 are presented in Figure 4 and Table 5.

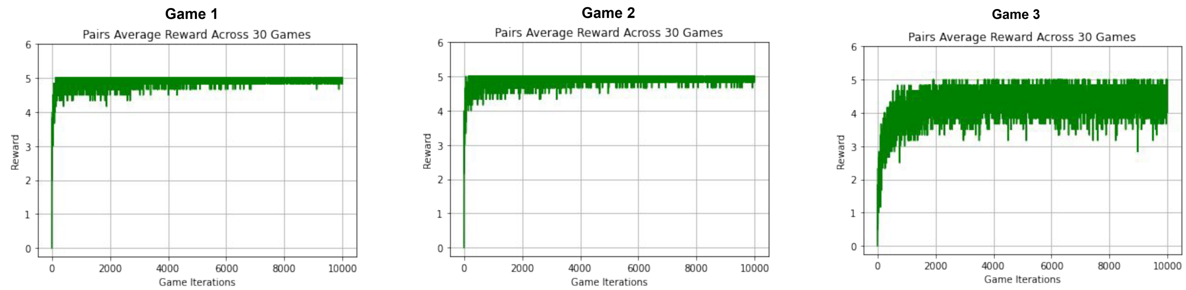

Fig. 4: This figure displays the results of different games for scenario 4.

Scenario 5 demonstrates an average agreement of approximately 98.8 percent across all games. The plots exhibit a convergence pattern similar to that of Scenario 4, but with a slight increase in the agreement percentage. This suggests that the modification implemented for the sender has indeed impacted the communication process positively. The averaged cumulative rewards, as before, refer to the rewards of the pairs in each round. All results for Scenario 5 are presented in Figure 5 and Table 5.

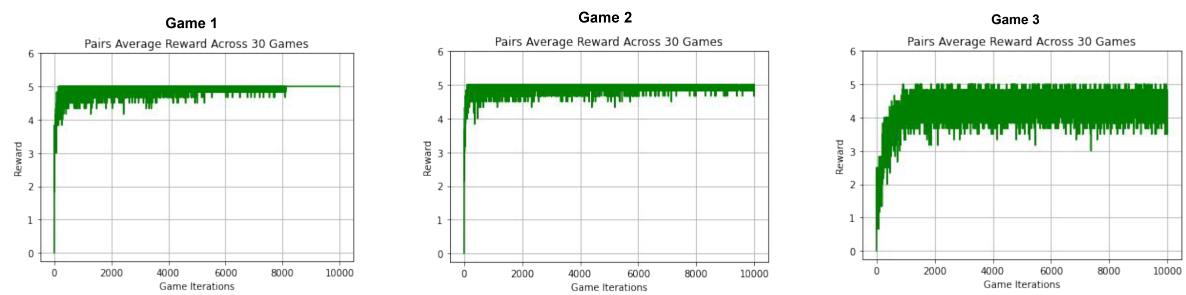

Fig. 5: This figure displays the results of different games for scenario 5.

Table 5: This table displays the results of different games for all scenarios.

		Games		
		Game 1	Game 2	Game 3
Scenario 1	Sender Signals	[['w2', 0, 0.99], ['w3', 1, 1]]	[['w0', 0, 0.99], ['w1', 1, 1]]	[['w10', 0, 1]]
	Receiver Signals	[['w2', 0, 0.99], ['w3', 1, 1]]	[['w0', 1, 0.99], ['w1', 0, 1]]	[['w10', 0, 1]]
	Agreement percentage	100%	100%	53.3%
Scenario 2	Sender Signals	[['w0', 0, 0.99], ['w4', 1, 1]]	[['w6', 1, 0.99], ['w9', 0, 1]]	[['w0', 0, 1]]
	Receiver Signals	[['w0', 0, 0.99], ['w4', 1, 1]]	[['w6', 0, 0.99], ['w9', 1, 1]]	[['w0', 0, 1]]
	Agreement percentage	100%	100%	56.6%
Scenario 3	Sender Signals	[['w9', 0, 0, 1], ['w13', 1, 1, 0.96]]	[['w0', 0, 0, 1], ['w2', 1, 1, 0.65]]	[['w6', 0, 1, 1]]
	Receiver Signals	[['w9', 0, 0, 1], ['w13', 1, 1, 0.96]]	[['w0', 1, 1, 1], ['w2', 0, 0, 0.65]]	[['w6', 0, 1, 1]]
	Agreement percentage	86.6%	86.6%	100%
Scenario 4	Sender Signals	[['w1', 0, 0, 0.99], ['w0', 1, 1, 1]]	[['w2', 1,1, 0.88], ['w3', 0, 0, 1]]	[['w0', 0, 1, 1]]
	Receiver Signals	[['w1', 0, 0, 0.99], ['w0', 1, 1, 1]]	[['w2', 0, 0, 0.88], ['w3', 1, 1, 1]]	[['w0', 0, 1, 1]]
	Agreement percentage	93.3%	83.3%	100%
Scenario 5	Sender Signals	[['w12', 0, 0, 1], ['w15', 1, 1, 1]]	[['w3', 0, 0, 1], ['w8', 1, 1, 0.9]]	[['w13', 0, 1, 1]]
	Receiver Signals	[['w12', 0, 0, 1], ['w15', 1, 1, 1]]	[['w3', 1, 1, 0.99], ['w8', 0, 1, 1]]	[['w13', 0, 1, 1]]
	Agreement percentage	100%	96.6%	100%

5 Evaluation

It is evident from the results that Scenario 4 exhibits a slightly higher agreement percentage compared to Scenario 3 for a population size of two. Scenarios 1 and 2 show the highest agreement percentages, but this is primarily due to the one-sided communication and fewer random variables involved in signal determination. The exploration rate has a noticeable impact on payoff convergence and can vary depending on the reward distribution schemes. Comparing the agreement percentages becomes more interesting when testing Scenarios 3, 4, and 5 with larger populations and reciprocal communication. For small groups with fewer than 10 agents, all scenarios demonstrate relatively high agreement percentages across all games with 50,000 iterations, with Scenario 3 having a slightly lower average agreement percentage. However, as the number of agents increases, Scenario 5 consistently achieves the highest agreement percentage, as shown in Figure 6. For a population size of 20, Scenario 3 exhibits a significantly low agreement percentage. This is attributed to the fact that all agents simultaneously adjust their signals, making them more susceptible to miscommunication. In Scenario 4, the selected pairs efficiently adjust their signals, but the agreement percentage suggests that the agents require more information to converge to a mutual semantic. Thus, the agreement percentage notably increases when senders are enabled to observe and incorporate the receiver's actions in finalizing their signals, holding true for all game variations. The displayed results only present the outcomes of Game 1 for simplicity.

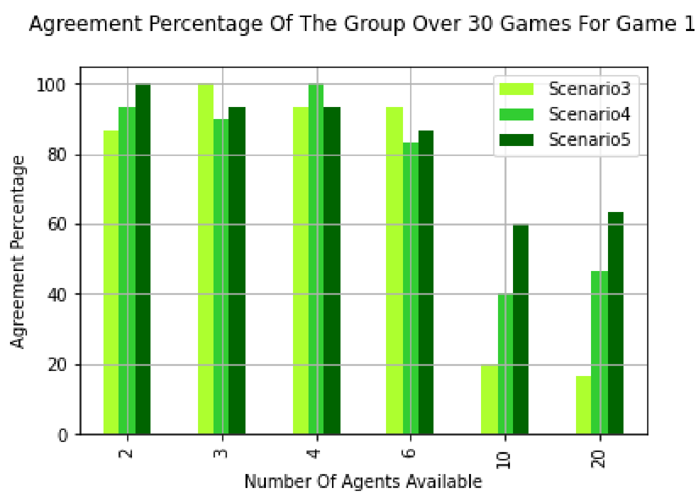

Fig. 6: This figure displays the agreement percentage of different scenarios and population of agents.

6 Conclusion

In this work, we proposed a communication scenario that enables intelligent agents to send and receive signals and take actions accordingly. We developed several games to analyze the success of communication and the performance of these agents in a simple scenario (Scenario 1) with one-sided communication. Subsequently, we implemented two learning methods, namely the epsilon-greedy approach, to optimize the confidence level of signals and the value of the roles. This optimization aimed to enhance communication and increase the agents' agreement percentage in optimized scenarios (Scenarios 3 and 4). The experimental results indicate that Scenarios 3 and 4 achieve a higher agreement percentage in communication compared to the baseline scenarios of 1 and 2. These findings demonstrate the effectiveness of the implemented learning methods in improving communication outcomes.

The research question of this paper was: *'To what extent do agents agree about the semantics of the emergent language in coordination games, and how can this agreement be further improved?'*. The analysis demonstrates that agents can achieve a notable agreement of almost 90% by learning the value of roles and the confidence of signals for a population size of two. This conclusion is supported by the figures displayed in Figures 1, 2, 3 and 4. Furthermore, we investigated the sub-research question of *'What factors contribute to reducing the time required to achieve higher agreement?'* by conducting experiments with different population sizes and allowing the sender to adjust their signals by observing the receiver's actions. The results show that in Scenario 5, where the sender observes the receiver's actions, the agreement percentage increases as the number of agents in the population increases. These findings align with the hypotheses stated in the introduction of the report.

7 Future Work

An interesting avenue for future research would be to investigate the impact of having more than two roles available for agents to choose from. Our scenarios focused on the effect of population size on the group agreement percentage, with only two available roles, which limited the variety of games. Therefore, expanding the agents' internal structure and conducting further experiments with multiple roles could yield interesting outcomes. As suggested by Kauhanen (2020a), individuals often adapt their language usage based on the linguistic behavior observed within their in-group while diverging from the practices of outgroups [12]. This phenomenon could be further explored by employing evolutionary algorithms on a population of our current agents to analyze the adaptation of generations over time. To achieve this, children that have evolved from different scenarios would need to work together to solve unseen games and adjust their existing signal bases with the newly formed groups. Lastly, conducting longer experiments with different reward distributions could also provide insights into the dynamics between game structure and the efficiency of agents in solving the games.

References

1. Spike, M., Stadler, K., Kirby, S., and Smith, K. J. (2017). Minimal Requirements for the Emergence of Learned Signaling. Cognitive Science, 41(3), 623–658. https://doi.org/10.1111/cogs.12351
2. Steels, L. (2012). Introduction. Self-organization and selection in cultural language evolution. In Advances in interaction studies (pp. 1–37). John Benjamins Publishing Company. https://doi.org/10.1075/ais.3.02ste
3. Nowak, M. A., Plotkin, J. B., and Krakauer, D. C. (1999). The Evolutionary Language Game. Journal of Theoretical Biology, 200(2), 147–162. https://doi.org/10.1006/jtbi.1999.0981
4. Kirby, S., and Hurford, J. R. (2002). The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model. In Springer eBooks (pp. 121–147). https://doi.org/10.1007/978-1-4471-0663-0_6
5. Tamariz M, Papa A, Cristea M, McGuigan N (2023) Context congruence: How associative learning modulates cultural evolution. PLOS ONE 18(4): e0282776. https://doi.org/10.1371/journal.pone.0282776
6. Kauhanen, H. (2020). Replicator–mutator dynamics of linguistic convergence and divergence. Royal Society Open Science. https://doi.org/10.1098/rsos.201682
7. Barrett, J. S., and VanDrunen, J. (2022). Language games and the emergence of discourse. Synthese, 200(2). https://doi.org/10.1007/s11229-022-03645-7
8. Izquierdo, L., and Sandholm, W. H. (2019). An introduction to ABED: Agent-based simulation of evolutionary game dynamics. Games and Economic Behavior, 118, 434–462. https://doi.org/10.1016/j.geb.2019.09.014
9. Peter Godfrey-Smith, Signals: Evolution, Learning, and Information, by Brian Skyrms., Mind, Volume 120, Issue 480, October 2011, Pages 1288–1297. https://doi.org/10.1093/mind/fzs002
10. Franke, M. (2008c). Semantic Meaning and Pragmatic Inference in Non-cooperative Conversation. In Lecture Notes in Computer Science (pp. 13–24). Springer Science+Business Media. https://doi.org/10.1007/978-3-642-14729-6_2
11. Lipowska, D. and Lipowski, A. (2018). Emergence of linguistic conventions in multi-agent reinforcement learning. PLOS ONE, 13(11), e0208095. https://doi.org/10.1371/journal.pone.0208095
12. Kauhanen, H. (2020a). Replicator–mutator dynamics of linguistic convergence and divergence. Royal Society Open Science. https://doi.org/10.1098/rsos.201682