CHALMERS
UNIVERSITY OF TECHNOLOGY

UNIVERSITY OF GOTHENBURG

Multi-agent Communication via Reinforcement Learning in Social Networks

Masters thesis in Computer science and engineering

ZHITAO LIANG
WANQIU WANG

Department of Computer Science and Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2024

Master's thesis 2024

Multi-agent Communication via Reinforcement Learning in Social Networks

Masters thesis in Computer science and engineering

Zhitao Liang Wanqiu Wang

UNIVERSITY OF
GOTHENBURG

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Computer Science and Engineering
Chalmers University of Technology
University of Gothenburg
Gothenburg, Sweden 2024

Multi-agent Communication via Reinforcement Learning in Social Networks
Masters thesis in Computer science and engineering
ZHITAO LIANG WANQIU WANG

Supervisor: Jonathan David Thomas, Computer science and engineering
Supervisor: Emil Carlsson, Computer science and engineering
Examiner: Devdatt Dubhashi, Computer science and engineering

Master's Thesis 2024
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2024

Multi-agent Communication via Reinforcement Learning in Social Networks
Masters thesis in Computer science and engineering
ZHITAO LIANG and WANQIU WANG
Department of Computer Science and Engineering
Chalmers University of Technology and University of Gothenburg

Abstract

This thesis investigates the use of multi-agent reinforcement learning (MARL) to explore emergent communications of artificial agents in social networks. The main goal is understanding how agents develop shared communication protocols to perform collaborative tasks in complex environments. Using the World Color Survey (WCS) dataset, we implement a speaker-listener model in which an agent learns to name colors, providing a framework for observing the formation of communication strategies. In contrast to existing work, we utilize a shared neural network for both speaker's and listener's functions, which promotes equivalence in language use between agents and supports consistent communication. Extending the model to multiple agents, we studied how social network structure affects emergency communication, finding that denser networks produce more consistent language while sparser networks allow for greater diversity. The introduction of new agents and different levels of interaction between communities also affects language evolution, with newly generated languages found to be more similar to more populous collectives. However, the scale of our research could be improved. In future work, investigating larger populations of agents would be beneficial for better understanding scalability and refining our findings. Additionally, we could explore other communication modes, such as one-to-many or many-to-one interactions, to gain a more comprehensive understanding of emergent communication in artificial systems.

Keywords: multi-agent reinforcement learning, emergent communication, social networks, color naming game, language evolution, World Color Survey dataset, thesis..

Acknowledgements

We would like to express our deepest gratitude to our supervisors, Jonathan David Thomas and Emil Carlsson, for their invaluable guidance and support in shaping this research project. We also sincerely thank our examiner, Devdatt Dubhashi, for his constructive feedback. Besides, studying at Chalmers University of Technology, particularly in Data Science and AI, has provided us with an incredible opportunity for growth and learning.

Zhitao Liang & Wanqiu Wang, Gothenburg, 2024-05-31

Contents

List of Figures

1

Introduction

1.1 Background

Communication through language is a crucial part of conducting large-scale coordination and sharing knowledge and experience among individuals. As different languages were formed in different human populations along with diverse cultures and societies, it's natural to consider how emergent languages develop in multi-agent artificial systems when moving forward to an era of AI.

The question of how different social structures and network typologies affect human language and communication systems has been a key topic of linguistic and cognitive science research for many years. Labov et al. stated that social structure shapes language structure [1]. Granovetter [2] found that sparser connections between members of a community promote more language variety while less stability, and more potential to change. In contrast, denser communities maintain higher consistency and conservative forms, as shown in the study by Milroy and Milroy, 1985 [3].

Raviv et al. [4] designed and conducted experiments to investigate the role of social structures in emergent human communication. They organized human participants into different types of social networks and asked them to play a communication game through which language patterns are formed in different social structures. Unexpectedly, their results didn't reveal any effect of social structures on language patterns in terms of communicative success, stability, and convergence.

Meanwhile, as shown by the study conducted by Ibrahim et al. [5], Multi-agent systems with different network typologies are applicable to specific types of cooperation tasks and scenarios. It's urgent to understand and explore how emergent language and communication patterns are affected by connection structures among agents in multi-agent artificial systems. In this project, we focus on emergent language patterns developed by artificial agents in different social dynamics in an artificial communication protocol via reinforcement learning.

Most existing work studying emergent language systems in artificial agents sets up experiments in which communication is restricted to interactions between two independent agents(for example, Lazaridou et al., 2017 [6], Bouchacourt and Baroni, 2018 [7], Kagebĺack et al., 2020 [8], and Carlsson et al., 2024 [9]). These studies adopted the speaker and listener communication model, in which language patterns are developed by asking the speaker to name an object and the listener to guess an object based on the name. The object could be color [8], images of dog or cat [6]. However, language formation through communication to achieve a cooperation target usually occurs in a larger community where multiple agents are able to interact with each other.

Thomas et al. [10] presented a setting where multiple pairs of speakers and listeners are allowed to develop potentially unique languages. However, in their implementation, speakers and listeners are separate agents learning and maintaining their own emergent communication patterns. This design is unlike the realistic scenario where each human in a community could perform two roles, both speaker and listener, learn and earn experience from the feedback of both speaking tasks and listening tasks.

Therefore, we aim to build a framework of a multi-agent communication system to support the following properties: (1)agents communicate with each other, (2)each agent can speak and listen, and (3)each agent forms its own language patterns from communication including speaking and listening experience. Based on this framework, we study agent-level communication and emergent linguistic behavior, exploring the conditions under which convergent and symmetric communication arises within a community. Furthermore, we analyze how the structures of these communities - specifically, the patterns of agent connectionsaffect the processes of convergence and language learning.

In addition to emergent languages in agent-level communication, contact linguistics in community-level interactions and mergers is also a natural and common phenomenon along the dynamics of social structures. Thus, it's an interesting topic in language and cognitive science. Myers-Scotton [11] studied contact linguistics and found that grammatical convergence occurs when two distinct languages come into contact with each other. She also addresses language attrition, where a less dominant language gradually loses its speakers and features due to the influence of a more dominant language. As computational models and AI have been increasingly used in the study of sociology, a need to model community-level creativity and evolution arises. It's pertinent to consider the emergent and evolutional phenomena in language when interactions occur between two separate groups of artificial agents. Thomas et al. [10] demonstrated that when an agent encounters a speaker of a different language, a period of adaptation is required before efficient communication can occur, which leads to the emergence of a new language and the gradual forgetting of the original language. What would happen if two groups of agents equipping different language patterns encounter? Several significant research questions arise: to what extent could the agents among two communities communicate with each other; could they convergent in new language patterns; how does the convergence and language evolution affected by the number of connections between two communities; what language patterns would a newborn agent learn when it connects to both communities; which original language would the new language be more similar to; would one language dominate another? To address these questions, we extend the agent-level framework to support community-level communication, and we examine the changes in language patterns of agents when they encounter and interact with agents from other communities.

1.2 Contribution

In this work, we successfully established a model in which agents are neural networks. They interact in a communication game - specifically, a color naming game, and learn experience and language patterns via multi-agent reinforcement learning. Meanwhile, we visualize the language patterns changes, and design and implement metrics to evaluate language patterns in terms of communication accuracy, quality, and consistency in this framework. In the study of agent-level communication, we demonstrate that although communities with all different typologies could end up with convergent language patterns, there are some significant effects of social structures on emergent language: (1) higher consistent and simpler language would be emergent in denser community, while the communication among the denser community is more efficient and accurate; (2) language complexity could increase the communication precision until a threshold is reached; (3) two closer agents tend to form more similar language patterns, but the difference becomes increasingly insignificant when the distance between two agents is larger. On the other hand, our work also presents some meritorious findings in emergent community-level communication: (1) when more connections are established, the agents within two communities are more likely to converge in common language patterns. (2) a newborn agent connected to both communities could learn and merge both distinct language patterns effectively. (3) the merged language patterns are more similar to the original language patterns of the larger population.

1.3 Limitation

At the same time, there are some limitations in this work:

1. **Population Size:** Our framework was tested on a limited population size, which may not accurately reflect the dynamics of larger social populations.
2. **Experimental Efficiency:** The current implementation requires about 30 minutes for every 80,000 communication, necessitating improvements for larger-scale experiments.
3. **Communication Types:** We focused solely on one-to-one speaker-listener pair communication and did not explore multiple-listener or multiple-speaker cooperation.
4. **Comparison Framework:** While visualizing policy changes and color partitioning, we lack a comparison framework with human language color partitioning patterns.
5. **More New Agents:** Our community communication experiments introduced only one new agent. Future studies should consider the impact of multiple new agents on language fusion and communication network efficiency.

1.4 Overview

This thesis is structured as follows:

In chapter 2, the theoretical background and concepts involved in this thesis are introduced. This includes discussions on the algorithms of reinforcement learning, the definition of diverse social network structures, the introduction of various computational frameworks for communication modeling and emergent language, and the theory of color partitioning.

In chapter 3, the relevant models and algorithms used in this article are introduced. This includes the data set used, the speaker-listener model, the color naming game, and metrics for evaluating the model. Futhermore, we describe the details of the experimental designs.

In chapter 4, the details of numerical and visualized results from the experiments including agent-level communication and intra-community communication have been shown and analyzed.

In chapter 5, conclusions from the results have been drown. Moreover, we discussed some limitations and possible extensions of our work.

In chapter 6, we summarize the entire thesis, providing a concise overview of our work.

2

Theory

In the chapter, the concepts and theories will be described thoroughly. First, reinforcement learning is used as the fundamental computational tool in the framework of the multi-agent communication model. In section 2.1, we started to introduce the concepts of the Deep Reinforcement Learning (DRL) algorithm, followed by the description of the Policy Gradient, REINFORCE algorithm, and Multi-Agent Reinforcement Learning (MARL). Then, sections 2.2 and 2.3 present the background and design of social network structures and the color naming, partitioning, and communication used throughout the work, respectively.

2.1 Reinforcement Learning

Throughout our project, the objective is to design and implement a framework that could support communication among multiple agents and create language patterns. Aiming at this goal, Reinforcement Learning, which is a computational approach to learning from interaction [12], is used. There are three fundamental machine learning paradigms, including supervised, unsupervised learning and Reinforcement Learning.

In supervised learning, the model is trained from input and labeled output pairs. The parameters of the model would be updated by computing the loss, which is the difference between the prediction of input and the corresponding ground truth output. This approach has limitations to learning from communication because it requires pre-labeled input-output data and cannot adjust to environmental changes and real-time feedback.

In unsupervised learning, the model expects to learn the hidden pattern or structure of the unlabeled data. It is also not an ideal paradigm to adapt knowledge from iterative interactions. Reinforcement learning is an appropriate machine learning approach that embodies the learning process through interactions, as discussed by Williams et al. in 1992 [13]. With two main components, the environment and the agent, RL enables the agent to observe the environment and make decisions with the goal of maximizing its cumulative reward, rather than uncovering hidden patterns within data, as shown in Figure 2.1.

In Reinforcement Learning, the agent could interact with the environment. Based on the state of the environment, it takes a specific action. After that, the environment

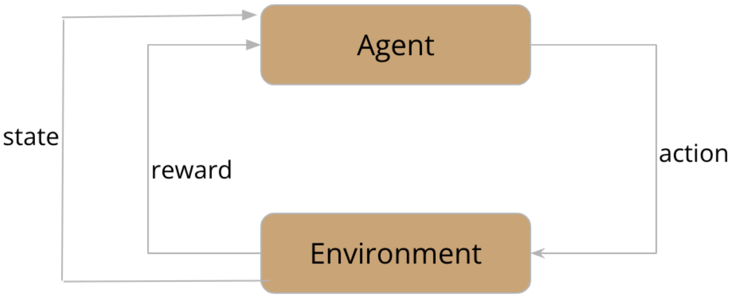

Figure 2.1: The agent-environment interaction in reinforcement learning.

would give feedback, which is usually a reward, to the agent, and the agent learns new experiences from the given feedback in each round of interaction. An agent makes decisions in two ways: exploration and exploitation. Exploration allows the agent to make random decisions in order to attempt a policy that hasn't been explored before and gain new experience. Meanwhile, exploitation means the agent could take the action that has been proven optimal by its previous experience. If an agent always selects the best action in its experience, it might miss the best reward in its unexplored action space and get stuck in the local optimal. On the other hand, if it always tries out new decisions, it might ignore the previous experience and loss a lot of rewards. By balancing exploration and exploitation, there is a trade-off that the RL algorithm should consider.

With the property of resembling how humans and animals learn from the natural environment, RL is a promising algorithm for artificial agents to simulate the language acquisition process, as Zaïem et, al. discussed in the review about learning to communicate in multi-agent reinforcement learning[14].

2.1.1 Deep Reinforcement Learning

Unlike classical reinforcement learning which uses tabular algorithms (for example, Q tables) to enumerate all possible actions in the current state, deep reinforcement learning makes use of multi-layer neural networks to replace the tabular methods for high-dimensional problems as the loop shown in Figure 2.2. When applying neural networks in reinforcement learning, the input of the NN is the observations of the environment obtained by the agent at each iteration step. The network is used as the value function or policy function, with nonlinear transformations applied to the input via weighted connections within each layer. Going through the complex transformations through the network, each network input is mapped to an output. As shown in Figure 2.3 The output at the network's last layer represents the probability distribution over the whole policy space at the current state, from which one action is sampled and conducted by the agent. Then, the environment gives feedback or reward to the agent, which is used to update the parameters of the neural network. Through the training process, the network is able to learn a set of parameters, which are the weights of connection in each layer, in which way it learns how to represent different abstractions. The details of the policy function, reward function and backpropagation will be discussed in Section 2.1.2.

We aim to design a framework that supports agent learning from interactions to simulate the language acquisition process. In the simulation, both the observation space and the action space might be high dimensional and sequential. DRL provides advantages in handling large observation and action space and high-dimensional data.

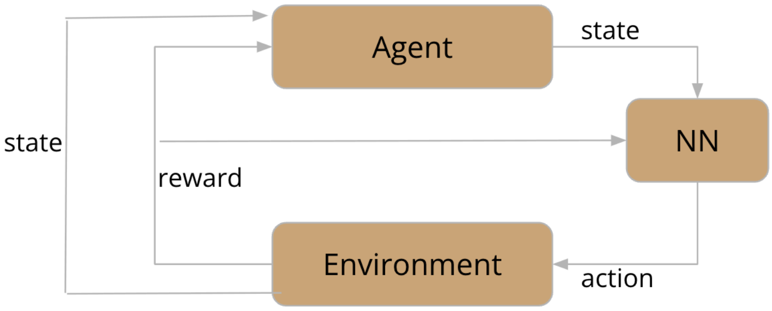

Figure 2.2: The agent-environment interaction in deep reinforcement learning.

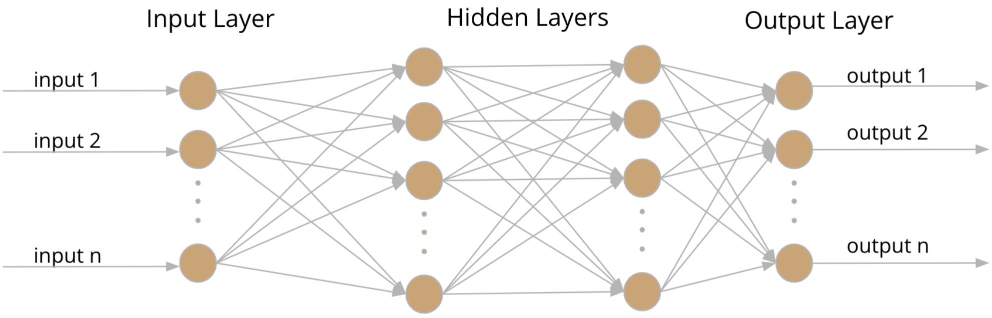

Figure 2.3: The architecture of a neural network.

2.1.2 Policy Gradient and REINFORCE algorithm

A policy is a strategy an agent uses to determine an action from the action set based on a given state of the environment. In other words, it maps the state to an action. Some reinforcement algorithms aim at optimizing the value function, which approximates the expected return value of being in a specific state or taking a specific action. Different from the value-based methods, policy gradient reinforcement learning, as introduced by Williams et al. in 1992 [13], models and optimizes the policies directly intending to improve the performance of policies by adjusting parameters of policy and increasing the return reward.

Here are some definitions and notations used in policy gradient methods:

- $s_t \in S_t$ - a set of possible states at time step t of one trajectory,
- $a_t \in A_t$ - a set of possible actions at time step t of one trajectory,

- $\pi_\theta(a|s)$ - the policy distribution function parameterized by θ, specifies the behavior strategy of one agent, in the form of the probability of an action in the condition of a given state,
- $Q^\pi(s,a)$ - return value or discounted future reward by taking action a in state s under policy $\pi(\theta)$,
- $d^\pi(s)$ - the state distribution under $\pi(\theta)$

To measure the performance of the policy, one objective function that gives the expected reward is defined:

$$J(\theta) = \sum_{s \in S} d^\pi(s) \sum_{a \in A} \pi_\theta(a|s) Q^\pi(s,a). \tag{2.1}$$

Therefore, the gradient of the objective function could be written as:

$$\begin{aligned}
\nabla_\theta J(\theta) &= \sum_{s \in S} d^\pi(s) \sum_{a \in A} Q^\pi(s,a) \nabla_\theta \pi_\theta(a|s) \\
&= \sum_{s \in S} d^\pi(s) \sum_{a \in A} \pi_\theta(a|s) Q^\pi(s,a) \frac{\nabla_\theta \pi_\theta(a|s)}{\pi_\theta(a|s)} \\
&= E_\pi[Q^\pi(s,a) \nabla_\theta ln \pi_\theta(a|s)]
\end{aligned} \tag{2.2}$$

As proposed by Sutton and McAllester et al. [15], the REINFORCE algorithm is a reinforcement method that utilizes policy gradient. It uses gradient ascent to update the parameter θ for policy $\pi(\theta)$ and maximize the expected reward under policy $\pi(\theta)$.

The algorithm collects the state, action, and reward at each time step to define a trajectory, as noted in equation 2.3.

$$\tau = (s_0, a_0, r_1, s_1, a_1, r_2, s_2, ..., s_T). \tag{2.3}$$

Each trajectory is used to compute the gradient of the objective reward function as defined in equation 2.1 and equation 2.2.

Then, the policy parameters θ are updated using gradient ascent to maximize the objective function in backpropagation, which is defined in equation 2.4.

$$\theta_{new} = \theta_{old} + \alpha \nabla_\theta J(\theta), \tag{2.4}$$

where α is the learning rate.

2.1.3 Multi-Agent Reinforcement Learning

In this project, to simulate emergent communication, including dissemination of language patterns and large-scale coordination within multiple individuals in society, we consider multi-agent systems, where each agent is considered an independent learner, as introduced by Tan in 1997 [16], rather than the multi-agent cooperation in a way with centralized training and decentralized execution as investigated by Sharma et al. [17], since the former is closer to the human learning procedure: each human maintains his or her own policies. Therefore, we adopted Multi-Agent Reinforcement Learning (MARL) with individual learners.

In single-agent reinforcement learning, only a single agent interacts with the environment and updates its behavior policy to maximize the reward within the environment. The state of the environment would only be affected by the action of a single agent. However, in multi-agent reinforcement learning, each of the multiple agents has its decision-making policy and keeps updating it, and each agent could take a specific action based on the state of the environment, which in turn changes the state, as the overview of MARL described by Buşoniu et al. [18]. The environment is no longer stationary, and the reward of a state-action pair could be different at different time steps depending on other agents' behaviors.

Ordinary reinforcement learning methods like deep reinforcement learning and policy gradient algorithms could also be used to address MARL problems [19]. Specific methodology depends on how the problem and the scenario are defined. For instance, if the agents are expected to learn from the interactions among each other, models such as multi-agent coordination communication games [6] are proposed to mimic this process. In these games, agents are placed in simple environments where they need to develop a language to coordinate and earn rewards from interactions. The details of multi-agent coordination communication games are introduced in section 2.3.

2.2 Social Network

Previous research, such as the work by Limor Raviv et al.[4], investigates how different social network structures impact the development of linguistic patterns. Their research involved conducting human experiments across various social network settings, demonstrating that different network configurations lead to distinct communication patterns. This work highlights the profound influence of social network structure on language evolution, showing that the type of network affects linguistic outcomes.

This thesis explores the connections and differences in developing new language formats on artificial agents experiments on different social network structures. The size of each social network will change in different experiments. In this part of the introduction, we temporarily assume that each network consists of eight agents. Each network differs in its degree of connectivity (i.e., how many other agents each agent interacts with). In addition, agents may have an edge connected to themselves, known as a self-loop, to model self-communication. This self-loop allows agents to process their own language and reinforces internal consistency in their language use.

Specifically, as shown in Figure 2.4, there are five types of social network structures: fully connected networks, small world networks, scale-free networks, ring networks, and random networks.

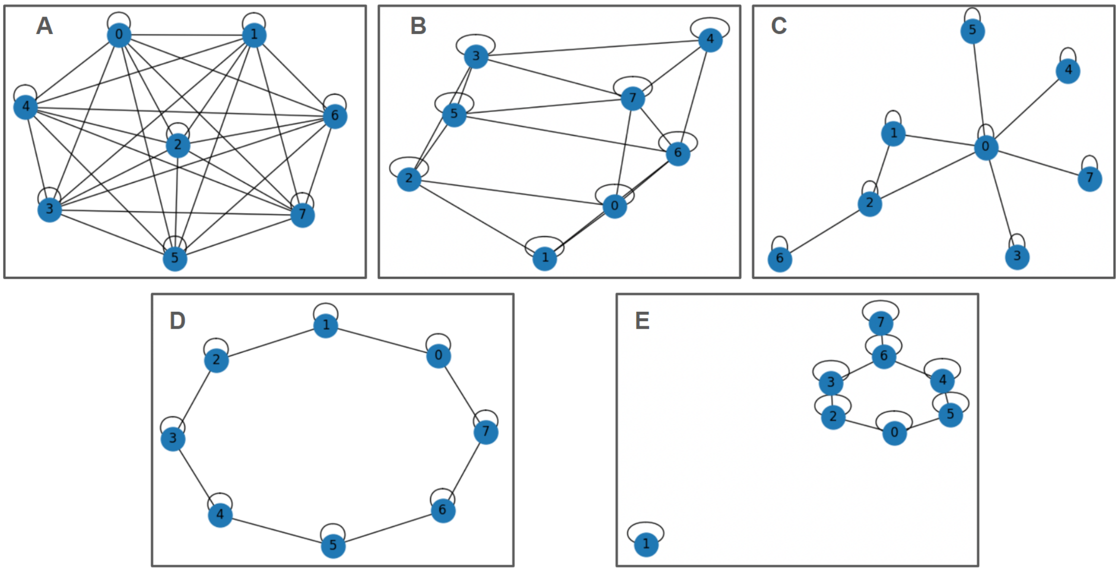

Figure 2.4: Types of social network structures. A) Fully-connected network, B) Small-world network, C) Scale-free network, D) Ring network, E)Random network

1. Fully connected network (Fig 2.4 A): In this network, all nodes can interact. So, every node has the same number of connections. This type of network is similar to early human societies, consisting of former tribes and small groups in which each member had direct opportunities for communication and contact [20]. In this social structure, information can be disseminated quickly and effectively. This network model is ideal for studying densely interactive phenomena such as disease spread or consensus building.

2. Small-world network (Fig 2.4 B): The average path length of this network is short, and most nodes are not directly connected, but nodes can also be connected through a few steps. For example, node 4 and node 5 are not directly connected but can be indirectly linked through nodes 3, 5, and 7. Small-world networks are much sparser than fully connected networks and only implement almost half of the possible connections. A well-known example of a small-world network is the Watts&Strogatz model, which generates a network with small-world properties by reconnecting the edges of a regular network [21].

3. Scale-free network (Fig 2.4 C): The distribution of node connections in this network follows a power law-distribution, meaning that there will be a few nodes (called "hub") with far more connections than the average [4] [22]. In contrast, most other nodes have only a small number of connections. For example, node 0 in the figure is the "hub" and interacts with almost all other nodes, while nodes 3, 4, 5, 6, and 7 are more isolated. This network model network widely exists in social networks such as the Internet and social media.

4. Ring network (Fig 2.4 D): Each node in this network is only connected to its two neighbor nodes on the left and right, forming a closed loop. Networks with this structure usually have a longer path length between two indirectly connected nodes, and the information propagation efficiency is not as efficient as other types

of networks. This type of network is often used to study local interactions and the fundamental properties of network protocols.

5. Random network (Fig 2.4 E) [23]: The edges between nodes in this network are randomly generated, and there are no other rules for connections. There may appear completely isolated nodes, such as node 1, which has no interaction with any other nodes.

An essential property of a network is the density of connections. A highly connected network means there are many direct connections between nodes in the network. In this network, any two nodes are almost always directly or indirectly connected over a short path to achieve fast communication. This type of network has a shorter average path length. Moreover, the degree (number of connections) of each node is high. This means that information can be spread quickly and efficiently throughout the network.

On the contrary, sparsely connected networks have fewer direct connections between nodes, and the degree (number of connections) of each node is usually low. This results in a longer average path length in the network, and communication between nodes may require indirect communication through more intermediate nodes, meaning information spreads slower.

For example, all nodes in a fully connected network are connected, and the degree of each node reaches the highest level. So, a fully connected network has highly dense connections. The degree of each node in the small-world network is almost half that of the fully connected network; the small-world network is sparser than the fully connected network. Furthermore, scale-free is also sparser than a fully connected network. Without considering random networks, the ring network is the sparsest among the above networks. Each node is only connected to two neighbor nodes, which is the node with the lowest number of connections.

Community integration is also an essential topic in network science, studying the structure and dynamics of social networks [24]. It describes the process of different communities combining or interacting in various ways in the network. Additionally, community integration affects the spread and formation of new languages.

For example, as Figure 2.5 shows, we can merge or interact with two independent communities by establishing a new connection or strengthening existing connections to achieve community integration. After being connected, different communities will begin to share information and may reach unified language and behavioral patterns.

2.3 Computational Framework for Multi-Agent Communication Modeling

Computational frameworks for multi-agent communication modeling are an essential aspect of emergent communication research. It defines the information and environment communicated between agents. First of all, communication is an essential aspect of framework design, so when designing games, we can either focus

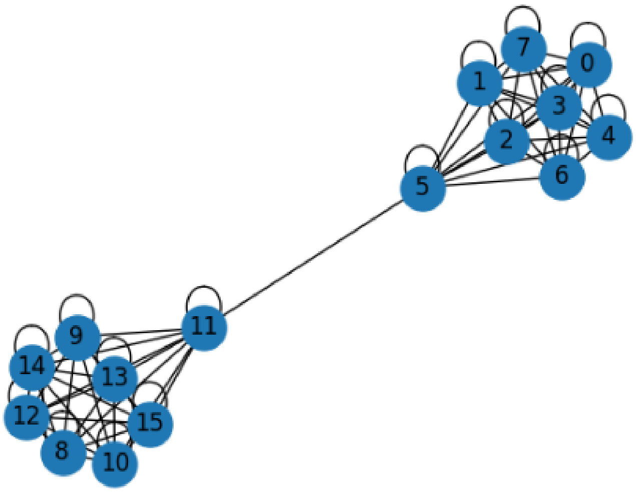

Figure 2.5: An example of community integration through a new connection

on the communication part or the part on achieving other goals through communication. For example, some studies mainly explore communication protocols between agents[25][26][27], while in other studies, the purpose of communication is only to help agents cooperate and make the agents more unified to achieve a common goal[28][29].

Next, we will introduce two critical frameworks in this field and use the combination of two frameworks in our experiments.

2.3.1 Referential Game

The first framework is called referential game, and one of the famous examples is discrimination referential game[30]. As Figure 2.6 shows a example referential game process:

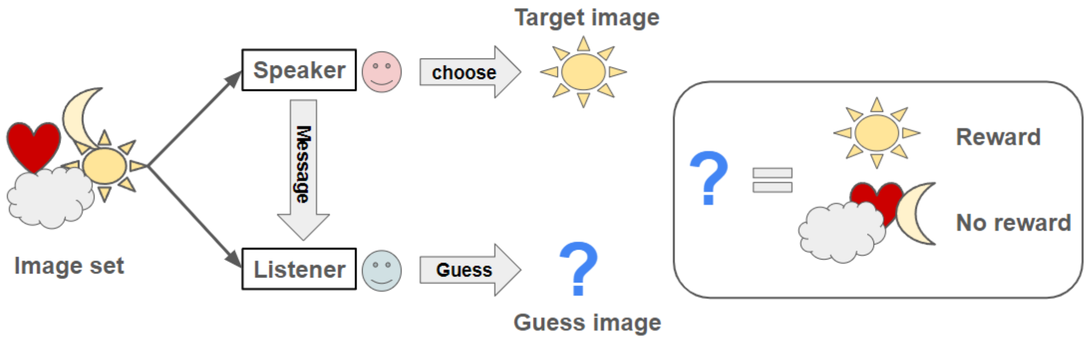

Figure 2.6: An example of the process of referential game.

1. There are two participants (agents), one set as the "speaker" and the other as the "listener."
2. Each agent receives a set of images.

3. The speaker randomly or specifically selects one of the images as the target (the sun) and generates a message describing the target image.
4. Based on this message, the listener guesses which image from the image set is the target image.
5. The final reward is determined based on whether the target image and the guess image are consistent; that is, the listener can accurately understand the information spoken by the speaker when successfully selecting the target image. If the listener guesses successfully, both participants receive a winning reward; otherwise, there is no reward.

There are many different types of referential games. For example, in the discrimination referential game introduced above, the listener must distinguish the target image from other images and then successfully select the target image. In addition, there is a generative referential game, where the listener generates new output results, such as re-creating an image based on the message sent by the speaker and then comparing them. The multi-step referential game is the same as the discrimination game, except that it is upgraded from single-step to multi-step communication. The listener can ask the speaker for more target information, and the game ends after the listener selection is completed. The multi-modal referential game refers to the different modalities obtained by the speaker and the listener. For example, one receives image information, and the other receives text information [31].

However, no matter what type of referential game it is, the primary research purpose is still the language produced by the speaker and listener during the communication process, and then conduct different experiments such as [25] [26] [32] [33].

2.3.2 Naming Game

As proposed by Steels and Loetzsch [34], the Naming Game is a simple language game experiment to explore the formation and evolution of human languages. In the Naming Game, the speaker names a given object according to its characteristic feature to draw the attention of the listener to this specific object. The object named by the speaker could be a color, a shape, or some other perceptually grounded category or relation.

Steels and Loetzsch [34] gave the definition of a simple scenario of the Naming Game: Assume there is a population of P agents, a set W containing objects in the game world and a vocabulary set V containing the words that could be used by the speaker to represent an object. A pair of agents is selected randomly from population P. One of them plays the role of speaker, and the other is the listener. The procedure of one round of the Naming Game is shown below.

Here, we define the concepts involved in the process:

- M - Vocabulary set: a series of messages that the speaker could use to describe the object. The message set could be finite or infinite, continuous or discrete, which depends on the goal of the game.

- $m_j \in M$ - The message selected by the speaker and sent to the listener.
- O - Object set: a series of objects that could be described by the speaker.
- $o_i \in O$ - The target object given to the speaker.
- $o_g \in O$ - The guess object selected by the listener according to m_j.

1. One object o_i is selected from the object set O as the target and given to the speaker.
2. The speaker names the individual target object o_i using one message m_j in the vocabulary set M.
3. The listener receives m_j, which is the name of o_i, from the speaker.
4. The listener searches its memory and experience and decides which object is associated with the name m_j given by the speaker.
5. The listener gave the speaker a signal about which object corresponds to m_j in its knowledge, for example, by pointing or ticking.
6. The speaker receives the listener's guess o_g and evaluates the game by comparing o_i and the object o_g pointed to by the listener. If two objects are the same one, then the game is successful. Otherwise, this round failed.
7. The speaker synchronizes the result of this round of games with the listener, and they both update their memory and experience.

This game requires each agent to retain a two-way memory associating the objects and the vocabulary used as their names, which allows the speaker to retrieve the name given an individual object and allows the listener to look up the object given the name. After a large number of rounds are carried out by the population, each individual in the population should have formed a stable and persistent memory, and the game could achieve steady communicative success as long as there are no other new objects introduced in this game.

Namine Game has been widely used in studies involving language acquisition in groups like social networks. Gao et al. [35] introduced an extensive variation based on the basic naming game, Naming Game in Groups (NGG), which provides a framework for Naming Game in typical network topologies including random-graph networks, small-world networks and scale-free networks, to examine the intra-group consensus achieved by Naming Game.

2.4 Emergent Language

Natural languages are so different from each other, which sparks researchers' interest in language emergence procedures. Emergent language refers to how common concepts and linguistics form within a period of time and how this could be used as a tool to support communication between human agents or artificial agents with a common understanding. Raviv et al. [4] investigated the emergent artificial languages developed by human agents in different networks in terms of their linguistic structures, communicative success, stability and convergence. However, their results didn't uncover any effect of network structures on these measures of emergent language. Their finding initializes our curiosity about how the connected structures of artificial agents affect the different properties of emergent language.

Some studies exploring the ability of artificial agents to develop emergent language and communication systems have also been conducted. These studies aim to explore the ability of artificial agents to form emergent communication patterns through interactions. Chaabounia et al. [36] used artificial neural networks in which a pair of speaker maintain their own respective neural network layers and play a discrimination game to develop a communication system whose distribution on the accuracy/complexity plane closely matches that of human languages. Thomas et al. [10] studied the adaptation of language when an agent encounters an agent with prelearned language. Their findings indicate that the emergence of a new language and the forgetting of the previous language occur after a period of adaptation.

2.5 Color Partitioning and Communication

Common Agreements in Color Naming Systems across Different Languages and Cultures.

Color naming and partitioning is an interesting topic in linguistic science and research. Lindsey et al. [37] studied the difference and universality of color partitioning across 110 different languages. Different languages vary in the number of color terms. Nonetheless, they have uncovered some common agreements in the color naming patterns that occur worldwide, and similar color naming systems are often seen in unrelated languages. All color terms within 110 languages tend to cluster into 11 distinct color categories, with glosses into some basic English color words: Red, Green, Blue, Grue, Yellow-Or-Orange, Brown, Pink, Purple, Black, Gray and White. Furthermore, They found the color naming systems across all 110 languages exhibit a dominant, clear RED region, and most of them display a clear Yellow-Or-Orange region, but the partitions of the cool region, including "green", "blue" and "purple" areas in English terms, vary considerably in different languages. As Zaslavsky and Kemp et al. [38] indicated in their study, warm colors have higher communication precision than cool colors, which could be explained by perceptual structure.

Informativeness and Simplicity Trade-off of Color Naming System in Human Languages.

As discussed by Regier et al. [39], informativeness and simplicity of languages could be considered as two competing principles in information theory. A language system that has a specific word for each distinguishable object could be regarded as highly informative, for instance, in the color domain, each separate color tile in the color map has a distinct term to represent it. A language system could be seen as simpler if the vocabulary set used by it is smaller, for example, if there is only one color term to represent all the color terms, the color naming system is regarded as maximally simple. In human language patterns, the trade-off between informativeness and simplicity could achieve optimal or near-optimal.

Kågeback and Carlsson et al. [8] proposed a computational framework that involves reinforcement learning to simulate the human language acquisition process and lead to optimal or near-optimal communication schemes. Chaabouni et al. [36] played a discrimination game with artificial neural networks trained with deep-learning techniques to develop communication systems in the color naming domain. Tucker et al. **??** proposed a method for training neural agents to encode inputs into discrete signals embedded in a continuous space, and they used multi-agent reinforcement learning settings in color reference games to evaluate.

3

Methods

3.1 Data

3.1.1 World Color Survey

In our project, the dataset that we used was the World Color Survey (WCS) (`www.icsi.berkeley.edu/wcs/`) [40]. The WCS research project was initiated by linguists at the University of California, Berkeley, in the late 1970s and resulted in a large and extensive color naming database. The naming data comes from adults in 110 different preindustrialized cultural societies around the world naming a set of standard color samples in their own languages. As Figure 3.1 shows, this standard color sample uses 330 color chips in the Munsell color system, covering almost the main range of visually distinguishable colors.

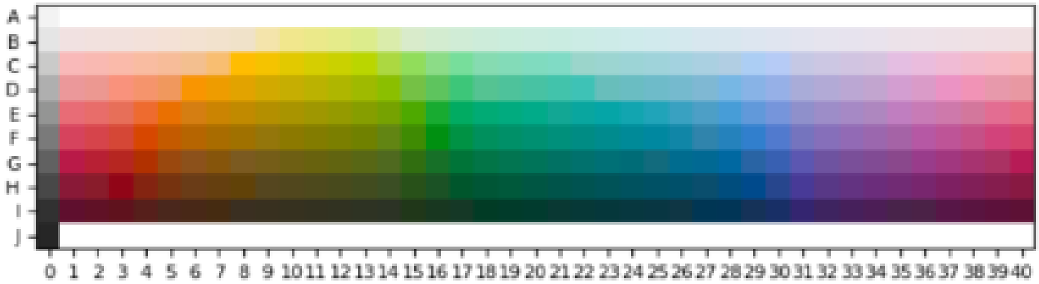

Figure 3.1: The color map of WCS.

This research project aims to record and analyze people's color naming habits in different cultural backgrounds. In Section 2.5, we also introduced the research by Lindsey et al. [37] on the differences and universality of color classification in 110 different languages. The results show that despite significant differences in the backgrounds of the experimental participants, humans' primary classification of colors is somewhat universal. For example, there are similarities in the perception and naming of basic color words such as "red," "blue," and "green." Still, the specific color boundaries and classification details will be different, such as whether green and blue are in the same category.

WCS data provides extensive data on color naming in natural language and is now widely used in research in linguistics and cognitive science. Therefore, we finally chose to use WSC data for model design. Based on this data, we study whether there is a certain universality or similarity between the color naming of artificial agents and the color naming of humans. Moreover, the differences in color boundaries and classifications between different artificial agents should be studied.

3.2 Communication Simulation

As the theory about social network structures shown in section 2.2, each agent is equivalent to one node in the social network. Our goal is to build a system that could support agents in a social network in communicating with each other. Therefore, we designed and implemented a framework to do communication simulation, including a speaker-listener model that indicates the sampling strategies to select a speaker-listener pair from the population network, the interaction between speaker and listener, and how the speaker and listener learn from the interaction. In a common way in the Multi-Agent Reinforcement Learning algorithm, all the agents share and maintain the same policy. However, in our framework, each agent is an independent learner and maintains its own decision-making policy, which resembles in a closer manner to real-life scenarios.

3.2.1 Sampling Strategy

As described in section 2.2, the environment of our model is a social network with different connectivity between each two agents within the population. If two agents connect to each other, they could interact with each other in the training process through which each agent forms their own language patterns. The connectivity between two agents reflects whether there is an edge between them. For Instance, as shown in Figure 3.2, looking into the agent (agent 0) in the center, there is a connected edge between agent 0 and agent 1, while no edge connects agent 1 and agent 2, which means agent 1 could communicate with agent 0, but couldn't interact with agent 2 in the training process.

In addition, the communication framework supports the self-loop of each agent. In other words, each agent could communicate with itself and update its own policy based on the feedback of self-communication. Therefore, there is also an edge connecting each agent and itself in the network shown in Figure 3.2.

In order to allow each agent to converge on a stable policy, a large number of iterations would be carried out. At each time step, the interaction would be conducted between a pair of agents. The sampling strategy to select the pair of agents in each iteration is introduced below:

1. Go through each agent in the population set.
2. For each agent A as the speaker in the population set, enumerate all its neighbors who have an edge connected to agent A as the listeners. Therefore, all possible pairs, (speaker, listener), in which the speaker is the agent A, are enumerated.

3. Combine all the possible pairs, and so far, we have the communication pair set, each element of which is a possible speaker and listener pair.
4. In each interaction, one speaker and listener pair is sampled from the communication pair set.

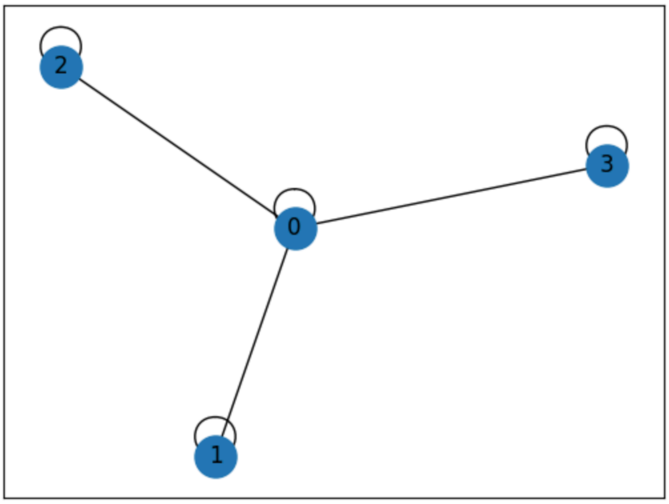

Figure 3.2: A Population Network with Four Agents.

3.2.2 Speaker-Listener Model

Some previous work introduced communication games to allow artificial agents to communicate with each other and learn from interactions. Kågeback et al. [8] use an architecture where the speaker and listener maintain separate speaking and listening policies in the task of color naming. Lazaridou et al. [6] described a framework to ask agents to name images. These research could provide a basic thought about establishing a communication framework between a pair of agents, based on which we are able to explore the effects of social dynamics on both agent-level communication and community-level communication. The overview of the Speaker-Listener Model is shown in Figure 3.3, in which a connected pair, agent 1 and agent 0, play the speaker and listener roles respectively.

However, in the existing work, the speaker and listener are separate agents, which is not that natural and different from the human manner in realistic scenarios where each human could speak and listen and earn experience from speaking and listening. Ohmer [41] proposed a way that is closer to human manner for a pair of agents: each agent could have speaking and listening behaviors. They suggested that the inference of the speaker mapping objects to messages and that of the listener mapping messages to objects are positively correlated. However, their theoretical framework does not include explicit implementations of speaker and listener roles within the model. Carlsson [42] implementated a framework combining speaking and listening abilities in one agent. However, each agent has two sets of parameters θ and ϕ in their framework: θ is used to map objects to messages when speaking while ϕ is used to map messages to objects when listening. We extended the speaker-listener model to allow every agent to be the speaker and listener simultaneously in a population of agents by implementing shared parameters in a neural network to complete speaking tasks and listening tasks as illustrated in Figure 3.4.

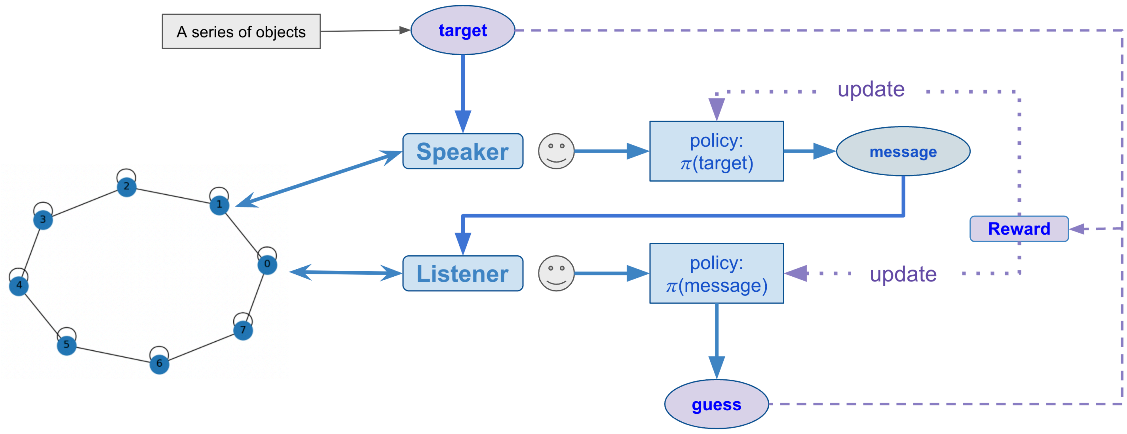

Figure 3.3: The Overview of Speaker-Listener Model.

Within the Speaker-Listener Model, the Deep Reinforcement Learning as introduced in Section 2.1.1 and Section 2.1.2, have been adopted. As shown in Figure 3.4 (B), each agent maintains one neural network with one input layer, one hidden layer, and one output layer. The neural network accepts the input, which is a combination of object and message representation, and outputs a score of y.

The details about how the neural network processes the input, how it makes the linear and nonlinear transformations and gains an output are shown in Figure 3.4 (C). For the speaking task, the inputs of the network are the target object combining all messages in the vocabulary set, and the multilayer perceptron would output one score for each of the inputs. All the output scores would go through a softmax function and become a distribution of messages. Then, one message would be sampled from the distribution by the speaker and sent to the listener. For the listener, the inputs are all the objects in the object set combined with the given message. Similar to the process on the speaker side, the distribution of objects would be calculated, and the listener samples one object as the guess object from the distribution.

Message Distribution For Speaking:

$$P(m|o_i) = S_{all\, m_k \in M}(f(o_i, m_k)), \tag{3.1}$$

Object Distribution For Listening:

$$P(o|m_j) = S_{all\, o_k \in O}(f(o_k, m_j)), \tag{3.2}$$

where $S_{all\, m_k \in M}$ denotes Softmax function over the message set, $S_{all\, o_k \in O}$ denotes Softmax function over the object set, and f(o, m) denotes the score output by the neural network.

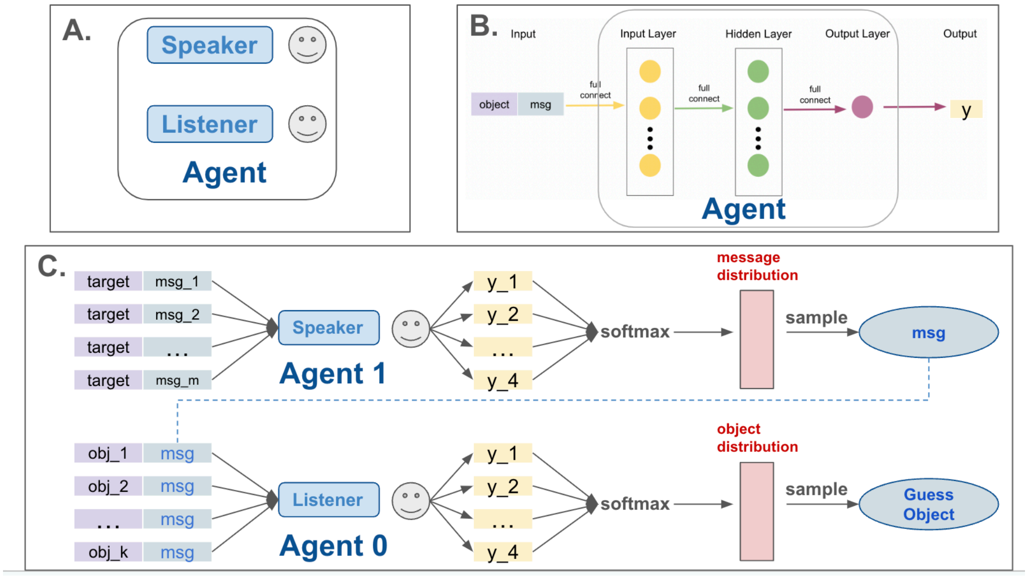

Figure 3.4: The Architecture of An Agent. A) Each agent in the population could act as both a speaker and a listener. B) The neural network within an agent. C) The calculation procedure by the neural network within the agent for Naming Game.

In backpropagation, the rewards calculated by comparing the target object and guess object would be used to update the parameters of the multi-layer neural networks within speaker and listener agents. We adopted the REINFORCE algorithm described in Section 2.1.2. The reward function and loss functions involved in this procedure are listed below.

Reward Function:

$$\begin{aligned} Reward &= 1,\ If\ the\ guess\ is\ the\ same\ object\ as\ the\ target; \\ &= 0,\ Otherwise. \end{aligned} \tag{3.3}$$

Speaker Loss Function:

$$E_{m \sim P(m|o_i)}[\nabla P(m|o_i) \log P(m|o_i) \cdot R] + \beta E[\text{entropy}(P(m|o_i))], \tag{3.4}$$

Listener Loss Function:

$$E_{o \sim P(o|m_j)}[\nabla P(o|m_j) \log P(o|m_j) \cdot R] + \beta E[\text{entropy}(P(o|m_j))], \tag{3.5}$$

Where R denotes the Reward.

3.2.3 Color Naming Game

Combining the Color Naming domain and the Speaker-Listener Model, we designed and implemented the framework of our experiments, the Color Naming Game.

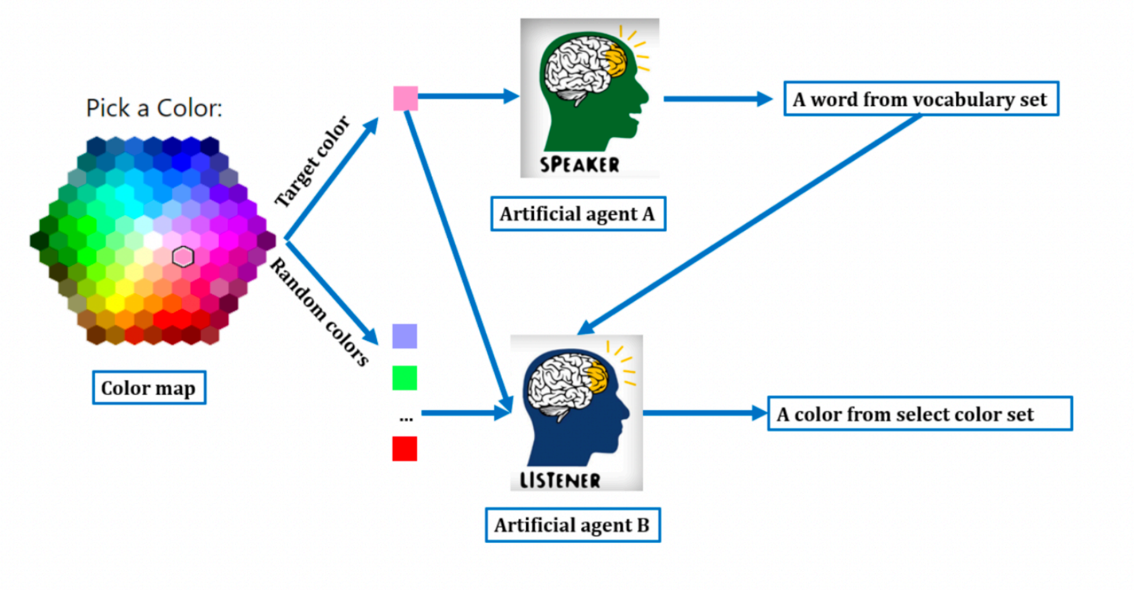

Figure 3.5: The Overview of Color Naming Game.

As shown in Figure 3.5, the objects defined in Section 3.2.2 are clarified as color tiles from the WCS color map introduced in Section 3.1, represented by three-dimension CIELAB data. The messages in the vocabulary set are represented as the one-hot representations of the indices. The actions of the speaker and listener are similar as described detailedly in Section 3.2.2.

3.3 Evaluation Metrics

In the evaluation of our model, we focus on the communication process and results and use some metrics to reflect important information, such as communication accuracy and language complexity. This section outlines some of the key metrics we use in our research.

Communication Accuracy

We consider the average communication reward value as a measure of communication accuracy. This metric quantifies the effectiveness of communication based on the results of interactions during the communication process. The higher the average reward value, the higher the success rate of communication; that is, the listener successfully guesses the same color as the target based on the message sent by the speaker; conversely, a lower reward value indicates poor communication effectiveness. This approach allowed us to clearly assess the impact of different experimental settings on communication effectiveness.

Language Complexity

In this thesis, we take two different approaches to assess agent language complexity:

1. **Conditional Entropy of Message Distribution:** This indicator is represented by the formula as follows:

$$H_a(M \mid O) = -\sum_i p_a(o_i) \sum_j p_a(m_j \mid o_i) \log p_a(m_j \mid o_i) \tag{3.6}$$

where the notations are defined as follows:

- $p_a(o_i)$ - the uniform probability of observation o_i according to the speaker agent a, equal to 1/330.
- $p_a(m_j \mid o_i)$ is the conditional probability of message m_j given the observation o_i according to the speaker agent a.

This metric evaluates the entropy of the message distribution for different input colors, reflecting the diversity of the agent's language once the speaker policies have converged. It measures how unpredictable the language use is after the policies have stabilized. Different studies use this metric differently, but typically, a low entropy value for a specific color indicates that the speaker consistently uses the same message to describe that color, showing a high degree of consistency [43]. In contrast, a higher entropy value suggests that the agent uses a broader array of vocabulary to describe a given color, demonstrating a richer and more varied linguistic output.

2. **Size of Used Vocabulary Set after Convergence:** The number of words used by the agents in a community after convergence of the emergent language system could serve as an indicator of language complexity. The size of vocabulary used by a language system could reflect the richness and diversity of language and communication between agents in this system. A larger vocabulary typically indicates a more complex language system that can describe the concepts with more detail. In the framework, the agents are provided a specific vocabulary set, and after a process of training, a convergent emergent language system occurs in the community. To evaluate the size of vocabulary in the convergent language system, we would run certain epochs of validation and collect all the words used in the validation stage, which could represent the used vocabulary set after convergence.

There is a correlation between these two metrics. Conditional entropy can indicate how unpredictable the language use is after the speaker policies have stabilized. However, if the system is highly certain about which words to use for each color(when reach stable low entropy), it does not provide information about the actual number of words used. On the other hand, vocabulary size directly measures the number of distinct words used by the agents after the language system has converged, but alone cannot capture the uncertainty or variability in language use, which conditional entropy provides. So, combining both metrics offers a more comprehensive understanding of language complexity. Together, they provide a holistic view of the communication language complex in artificial agents.

Color Partitioning

Color partitioning refers to mapping each of the 330 colors to the word with the highest probability $P(w \mid c)$, where w represents the word and c represents the color. Based on this division, we can get a picture as Figure 3.6 shows. The picture clearly shows the difference in color naming and classification of agents' language systems. To illustrate the visualization of color partitioning clusters, we collected the 3-dimensional cielab values of color terms in a cluster, sorted and got a median value of each dimension in each cluster, and then converted the cielab value to RGB representation, reflected as the color representation of each cluster in Figure 3.6. We also used different colors to represent different words respectively, as shown in the color of the edge for each cluster. In order to more intuitively quantify and compare the similarity of color naming and classification between the two language systems, we adopt the following indicators:

Figure 3.6: An example of a color partition map obtained after agent training.

1. **Rand Index:** Rand index similarity is a measure of similarity between two data clusters. In our project, we use the adjusted rand index [44] proposed by William M. Rand to compare the similarity between the color cluster maps generated after different agents communicate, that is, different agents' language systems. The degree of similarity that divides the color space. The higher the score of rand index similarity, the greater the similarity in color classification between agents, which also means that they use similar language systems to classify WCS colors. Specifically, the metric counts the number of occurrences of the same color in two agents color clusters.

2. **Well-formedness:** This metric is used to evaluate the quality of the divided color space [45]. Measure the similarity of colors within the same category and the color differences between different categories. In our implementation, by traversing each color cluster, calculating the similarity between different colors in the cluster, and calculating the difference between different colors in different clusters, the sum of similarity and difference is calculated as a well-formedness value. The higher the well-formedness value, the higher the quality of color space division.

Language Consistency

Language consistency is an essential indicator for measuring the similarity of language systems of different agents in social networks. In the color partitioning section, we introduced the comparison of the color classification of different agents. This only involves dividing similar colors into a cluster and comparing whether different clusters are identical. This obtains the similarity in the perception and classification of the same color by different agents. The language consistency within the network measures the consistency of agent communication and reflects their ability to maintain consistency in language use after communication. That is to say, for the same color, whether different agents use the same word to represent it, the agent can also guess the color with a high probability based on this word. In this project, we used two different metrics to calculate language consistency.

1. **Joint Distribution Consistency:** This metric evaluates the consistency of language use by comparing the joint probabilities of color-word pair selections by different agents; that is, it is used to analyze the consistency of selections by different agents in the entire color-word space. Specifically, what is first calculated is the joint probability of each agent choosing color c and getting word c. Then, the cosine similarity of the joint probability distributions between different agents is calculated. Mathematically, it is defined as:

$$\textbf{consistency}(a_i, a_j) = E\left[\sum_{m\in M}\sum_{o\in O} \cos\_sim\left(P_{a_i}(m,o), P_{a_j}(m,o)\right)\right] \tag{3.7}$$

   where $P(a_i(m,o))$ and $P(a_j(m,o))$ are the joint probabilities indicating the correlation of word m and color o for agent i and agent j. Higher value indicates higher similarity of language patterns between different agents, which means effective communication and a unified color naming method.

2. **Color Cluster Consistency:** This metric is derived from the color partitioning framework, which associates each color with the word with the highest probability. The consistency of each color is calculated by comparing the word with the highest probability assigned to that color by different agents. If the highest probability word in the message distribution of a color is the same in both agents, the consistency score for that color is 1; otherwise, it is 0. Overall consistency is the average of these scores for all colors. Mathematically, it is defined as:

$$\textbf{Color cluster consistency} = \frac{\sum_{c\in\text{Color}} \text{Consistency}(c)}{\text{Number of colors}} \tag{3.8}$$

   Among them, consistency(c) is 1 if the highest probability words in the message distribution of a given color c are the same and 0 otherwise. Higher values indicate greater linguistic consistency across agents.

3.4 Experiments

3.4.1 Emergent Color Naming Languages within One Community

3.4.1.1 Exp 1-1: Communication between a Pair of Agents

The communication between one pair of agents is a fundamental experiment in our study. There are two agents in this communication network, as shown in Figure 3.7, and they play the role of speaker and listener, respectively, in turn. In this experiment, we studied to address some problems: Could the artificial agents form and converge a language pattern in the color domain? How do the policies within one agent change and affect each other over a large number of communications? Do the color partitioning patterns formed in an artificial agent have some similar properties as those in human languages as described in Section 2.5?

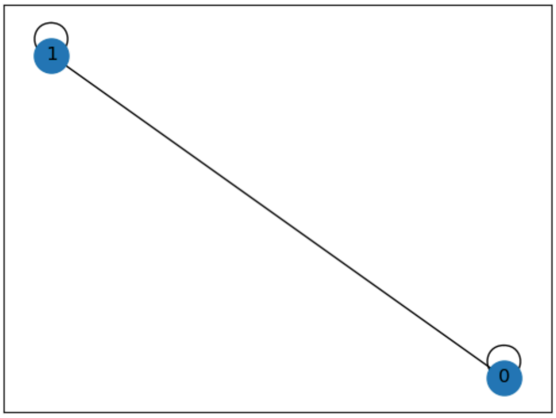

Figure 3.7: The Community of a Pair of Agents.

To analyze the research questions in our minds, we collected data about color partitioning maps, which show the policies, the policy change rates of the two agents, and the similarity of color clusters between the pair of agents during the training procedure.

3.4.1.2 Exp 1-2: Communication within Community with Four Agents

After we had a basic awareness of how the policies within different agents in one community affect each other during a series of communications, we moved forward our pace to study more complex communities. Starting with four agents, we conducted experiments on four different connectivity structures: one is a fully connected network, one is a ring network, one with a central agent, and one with a local cluster, as shown in Figure 3.8 (A), (B), (C), and (D) respectively.

In this experiment, we compared the policy changes of and the color partitioning patterns formed by the agents with different connectivity with other agents. In this way, we studied how connectivity affects the emergent languages in terms of color domain.

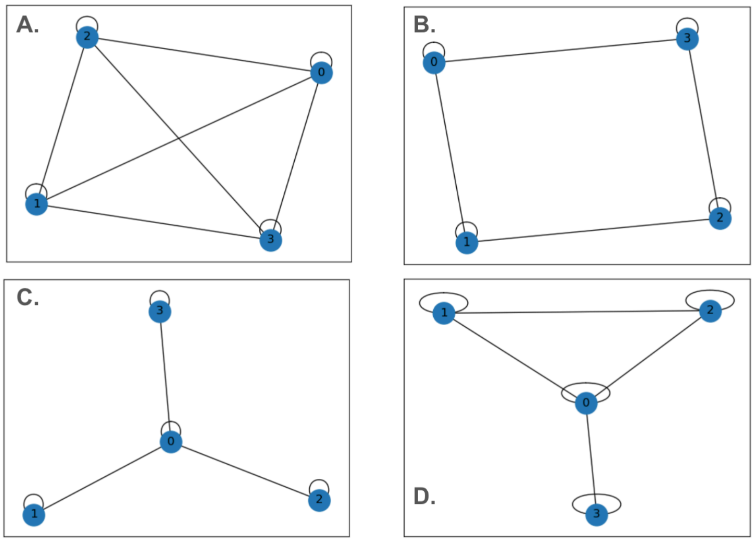

Figure 3.8: The Communities of Four Agents. A) A Fully Connected Network. B) A Ring Network. C) A Community with One Central Agent. D) A Community with A Local Cluster and An Agent Connected to One Agent within the Cluster.

3.4.1.3 Exp 1-3: Communication within Community with Eight Agents

We increased the complexity of the population to eight agents and studied the emergent language patterns within the communities in which they are organized as different structures, as shown in Section 2.2 and Figure 2.4

In this experiment, we use the size of the vocabulary, which is the number of given messages that a speaker could select, and use to represent objects and interact with a listener, and the average connectivity degree, which indicates the average closeness within one community, as two different variables, to study how these two properties affect the emergent language of color naming.

3.4.2 Communication between different communities

In the communication experiment between communities, we explore communication between different communities by establishing two different language communities (here, we use the fully connected social network as the community). We name the two language communities "Community A" and "Community B," each with its internal (intar-connections) and interaction (inter-connections) communication dynamics. Additionally, we name the agents that interact between two communities as bridge agents. Bridge agents act as intermediaries between the two communities, facilitating communication between them. The existence of bridge agents plays a crucial role in promoting language evolution between communities. The basic settings are shown in the Figure 3.9 below:

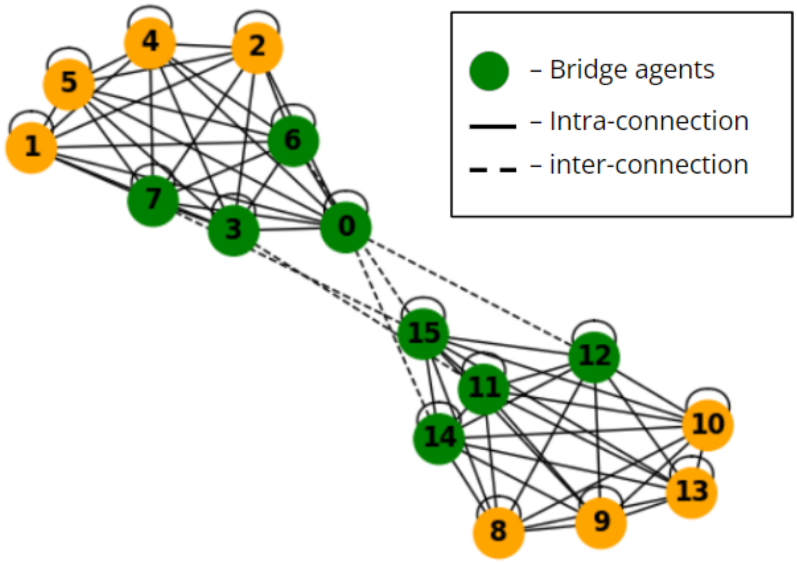

Figure 3.9: The basic setup of the inter-community communication.

3.4.2.1 Exp 2-1: The impact of varying degrees of interaction between communities on the formation of new languages.

This experiment focuses on exploring how different levels of interaction between the two communities affect the development of new language structures and shared vocabulary.

Settings and parameters:

First, we set up two communities in this experiment: "community A" and "community B". Each community comprises a fully connected network with a population of 8. The network for each community was pre-trained before the experiment. As mentioned earlier, there are internal connections (intra-connections) within communities, and interactions between communities are achieved through external connections (inter-connections). The communication link between communities A and B is represented by the dotted line connecting the bridge agents of the two communities, which is also the bridge for language integration between the two communities.

In order to quantify the interaction between communities, we introduce a new parameter, "Connect Num," which represents the number of inter-connections, which is the number of dashed lines shown in Figure 3.9. At the same time, as shown in Figure 3.10, we set up agent pairs called "inter-agent pairs" to represent all possible interactions between two communities, where one agent is from community A and the other is from community B. The connected bridge-agent pairs and connection edges in the figure are defined as "bridge-agent pairs" and "bridge-agent communication." The unconnected inter-agent pairs and connection edges are defined as "other-agent pairs" and "other-agent communication."

Communication dynamics:

By dynamically adjusting "Connect Num," we can observe the language spread and

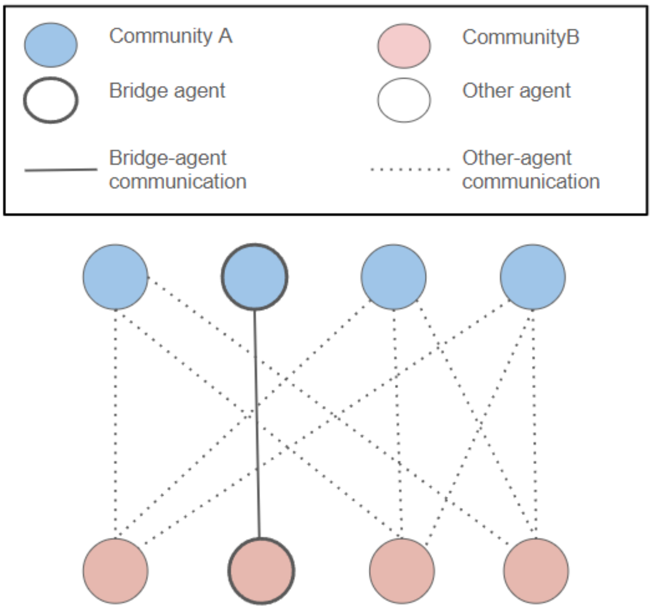

Figure 3.10: The setup of the bridge-agent communication and other-agent communication.

changes between the two communities. When conducting different "Connect Num" experiments, to ensure the consistency of the experiments, both communities A and B will be reset to the state after initial independent training. So, our specific experimental process for a certain "Connect Num" is shown in Figure 3.11 below:

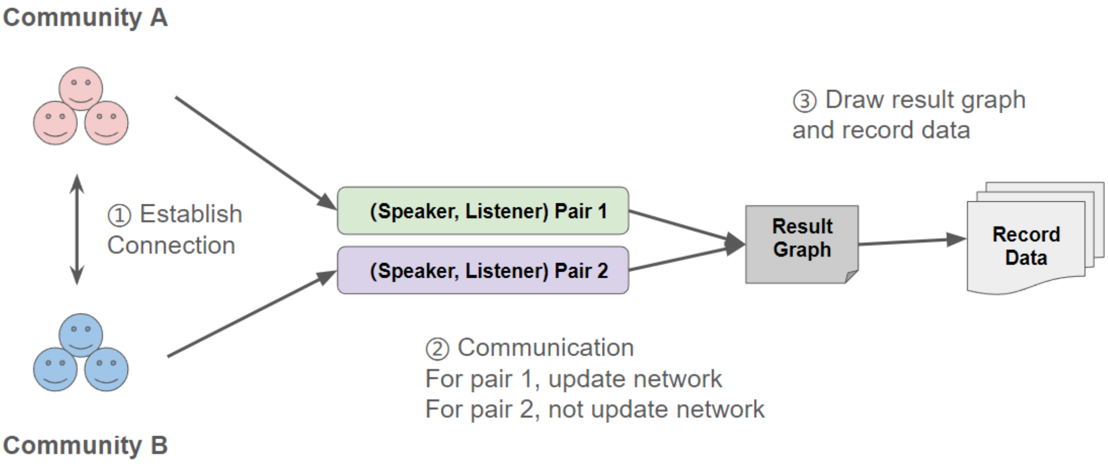

Figure 3.11: The experiment process for inter-community communication.

1. First, establish a connection between two communities by randomly but not repeatedly selecting "Connect Num" inter-agent pairs. As shown in Figure 3.10, the selected pairs are set to "Bridge communication pairs," and the unselected pairs are set to "Other communication pairs." Set the two agents in pairs as bridge agents. Add connections to the selected pairs in the network, representing inter-connection.

Item Randomly select a speaker-listener pair (in Figure pair 1) and then communicate. If the pair that we choose from is a "bridge-agent pair," then we will communicate with "other communication pairs" (in Figure pair 2) at the same time. Since "other-agent pairs" are not directly connected in the network, the communication here is only used to calculate the communication effect. It does not involve backpropagation updates of the network.

2. Based on the results after the experiment, draw the result graph and, record the data, and analyze the language fusion effect under different connection numbers.

3.4.2.2 Exp 2-2: The impact of adding a new agent (newborn) between communities on the formation of new languages.

In this experiment, we study the impact of introducing a new agent (called "newborn") between two pre-trained communities (Community A and Community B) on aspects of language communication and cultural integration.

Settings and parameters:

This experiment also consists of two fully connected networks representing community A and community B, each network consisting of eight agents. A new agent (newborn) is then introduced, which acts as an additional bridge between the two communities. The agents of the community that communicate with newborns are still named bridge agents. As shown in Figure 3.12, the two communities are connected through bridge agents (green nodes) and new agents (blue nodes). The communication between the new agent and the community is represented by a dotted line, which we define as the new-community connection.

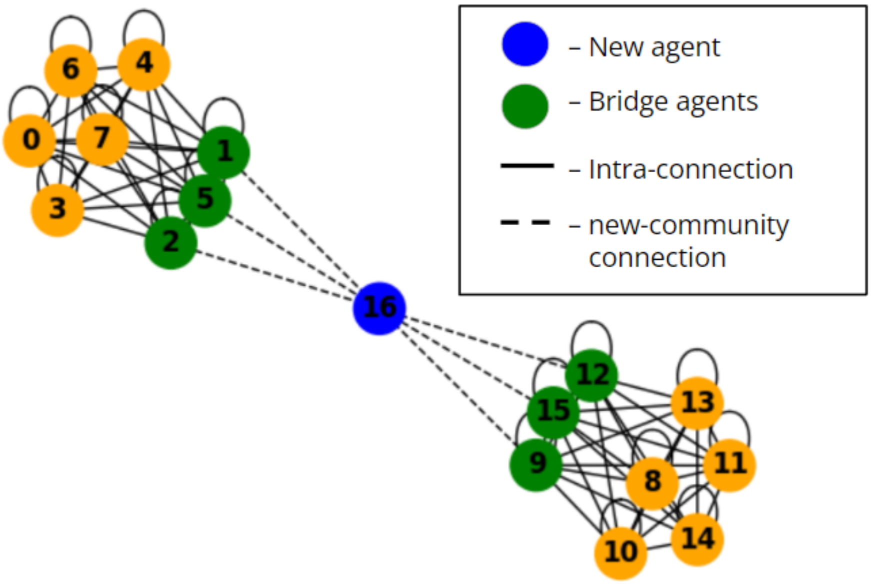

Figure 3.12: The setup of the inter-community communication, which adds a new agent between two communities.

We set a new parameter for this experiment. The new parameter is "New Connect Num," which represents the number of connections between the new agent and the two communities, i.e., the number of dashed lines in the figure. In addition, we continued the settings of "bridge agent" and "other agent pairs" in the previous experiment. As shown in Figure 3.13 below, we added "other new agent pairs" and "new agent pairs." At the same time, these connections are named "new-community communication" and "other new-community communication."

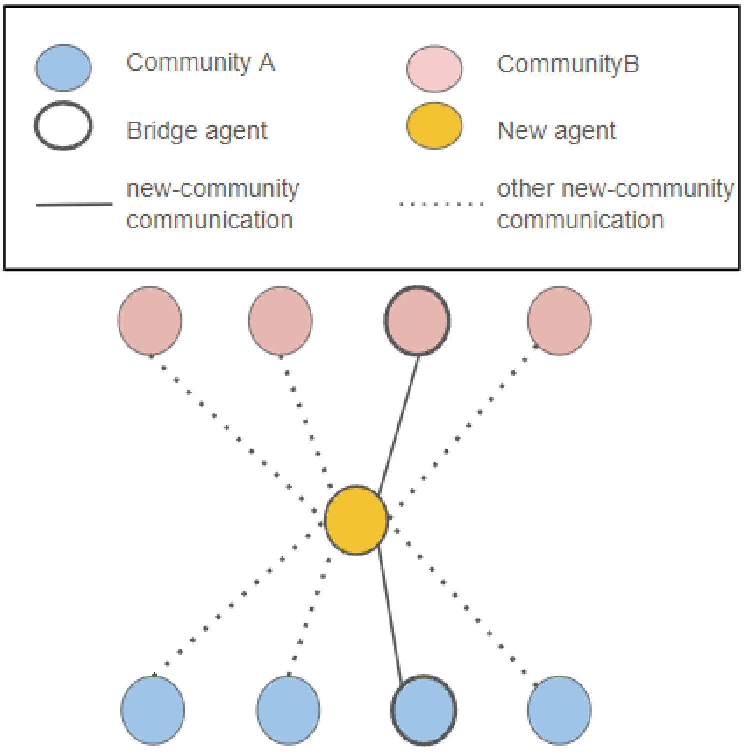

Figure 3.13: The setup of the new-community communication and other communication.

Communication dynamics:

The process of this experiment is similar to the previous experiment, which dynamically adjusts "New Connect Num" to observe the language spread and changes between the two communities and the new agent. After initial independent training before each experiment, we reset the community to the state. The specific flow chart of the experiment is shown in the Figure 3.14:

1. First, establish a connection in the network diagram: According to the settings of "New Connect Num, " select agents from the two communities and establish a connection with the new agent. For the parameter" New Connect Num," we distribute it equally to the two communities, randomly select agents from the two communities based on this number, and connect these agents to the new agent. As shown in Figure 3.13, the pair of the selected agents and the new agent is set to "new agent pairs." The unselected agent and the new agent together form "other new agent pairs." At the same time, all possible

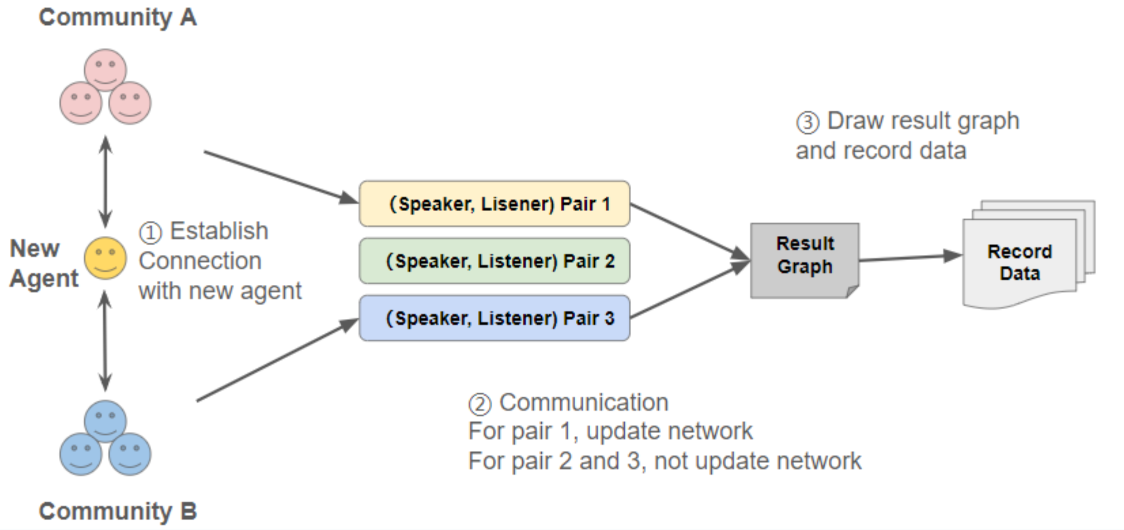

Figure 3.14: The experiment process for communication between the new agent and two communities.

communication pairs between two communities are set as "other agent pairs."

Item Randomly select a speaker-listener pair (in Figure, Pair 1) and communicate. If we select "new agent pairs," we will simultaneously conduct three communications of "other new agent pairs" (in Figure Pair 2) and "other agent pairs" (in Figure Pair 3). The communication here is also only used to evaluate the communication effect and does not involve network updates.

2. Based on the results of the experiment, draw the result graph, record the data, and analyze the language changes of agents and new agents in the community under different settings.

Although the final network diagram is the same in experiments with two different parameters, the communication process will be completely different.

3.4.2.3 Exp 2-3: The effect of the ratio of the population sizes of two communities on the formation of new languages.

In this experiment, we further explore the impact of the population size ratio between the two communities on the formation of new languages. We designed two different experimental programs: one program did not directly introduce new agents to compare the changes in different population sizes, and the other added new agents to study the impact of varying population sizes when introducing newborns as a communication bridge.

These two programs aim to observe how this change affects language integration and communication efficiency between the two communities by adjusting the number of community members.

Settings and parameters:

The experiment first set up community A as a fully connected network of 12 agents. A new parameter, population ratio, is introduced, which is calculated by dividing the number of people in community B by the number of people in community A. The population sizes of community B are set to 2, 4, 6, 8, 10, 12, 14, 16, and 18, respectively, thus forming different community size ratios.

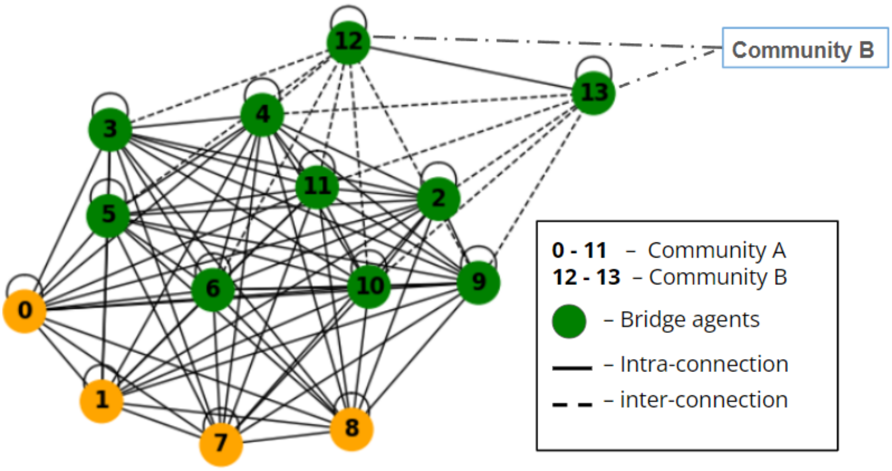

Figure 3.15: The setup of the experiment 2-3-1 (Without new agent), the population size of community B is 2.

Experiment 2-3-1 (without adding new agents):

As shown in Figure 3.15, two communities with a set population size are connected to explore how language communication between communities is affected only by changing the population size of community B without a new interactive agent. According to our hypothesis, the resulting language should be more similar to communities with larger populations.

Experiment 2-3-2 (adding new agents):

As shown in Figure 3.16, two communities are connected through a new agent to study how communities of different population sizes affect the evolution of language and communication dynamics when a new agent is added as a communication bridge. Based on our hypothesis, the new agent's language should be balanced between the two languages and tend to be more similar to the more populous community. With a new agent, the success rate of direct communication between the two communities will be lower than in experiment 2-3-1.

Communication dynamics:

The experiment adjusts the population ratio parameter by dynamically changing the number of agents in community B. The network of community A was reset to its initial independent training state before each experiment. The specific experimental process is shown in the following Figure 3.17:

1. In Experiment 2-3-1, the number of connections was dynamically adjusted based on the number of agents in community B and set to two-thirds of the

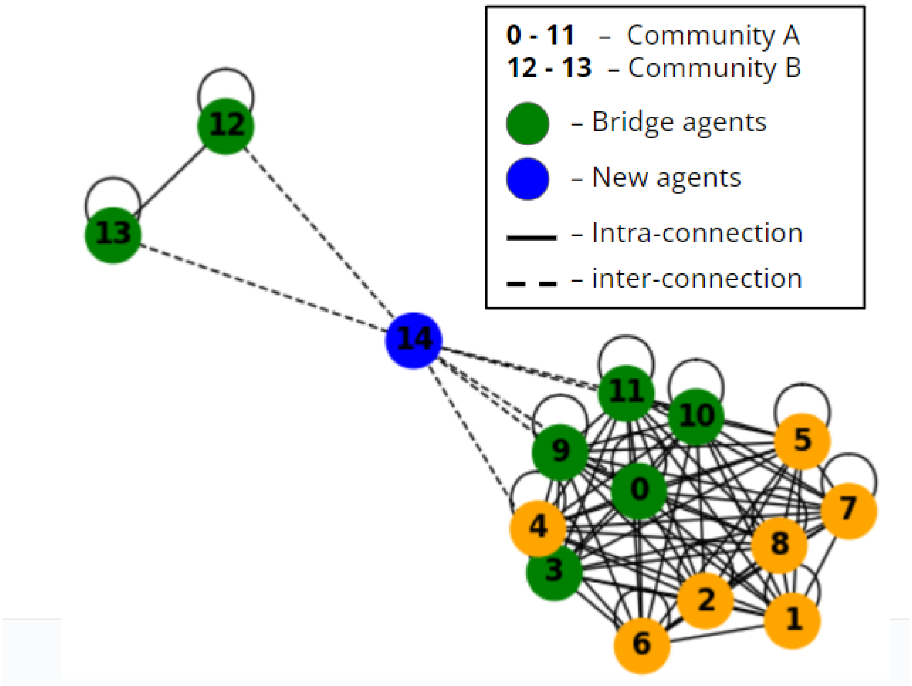

Figure 3.16: The setup of the experiment 2-3-2 (With new agent), the population size of community B is 2.

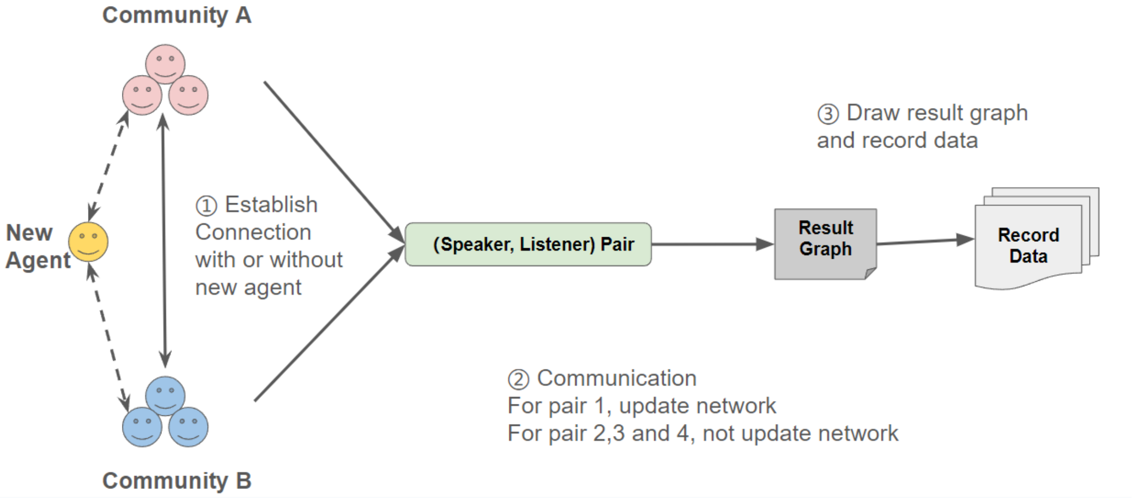

Figure 3.17: The experiment process for the communication of different community population sizes.

number of inter-agent pairs. Then, agent pairs are randomly selected from inter-agent pairs for connection based on the number of connections. In Experiment 2-3-2, a new agent was added, and the number of connections was dynamically adjusted based on the number of agents in community B. It was set to the sum of 12 and the number of agents in community B. That is, all agents in the two communities were connected to the new agent.

2. Randomly select a pair (speaker, listener) and then communicate.
3. Based on the results of the experiment, draw the result graph, record the data, and analyze the effect of different community population ratios on language integration. And the impact of adding a new agent on the formation of the new agent's language.

4

Results

4.1 Emergent Color Naming Languages within One Community

4.1.1 Exp 1-1: Communication between a Pair of Agents

Figure 4.1: The Color Partitioning Patterns Changes of Two Agents through Communication

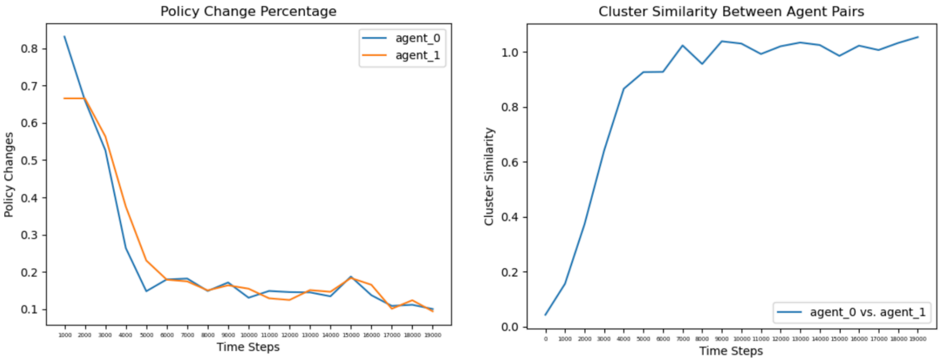

Figure 4.2: The Policy and Similarity Changes of Two Agents through Communication

Figure 4.1 shows how the policies of two agents in a community change in the training process. The color partitioning policy has been initiated differently for the two agents. Their policies affect each other, and the similarities have been increasing, as also shown in Figure 4.2 (B). In Figure 4.2 (A), we can observe that at the first 5000 training steps, the policies of both agents change dramatically. After that, the change slows down, and the policies become stable, which shows the two agents could develop stable, consistent, convergent language patterns through communication.

4.1.2 Exp 1-2: Communication within Community with Four Agents

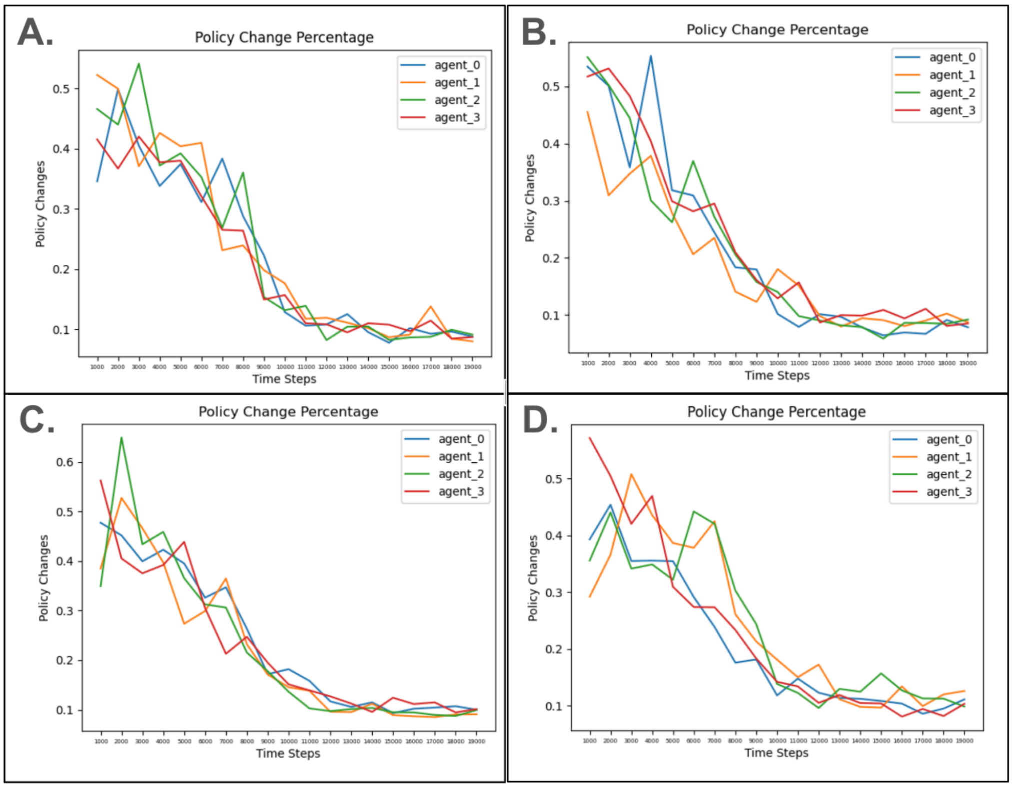

Figure 4.3: The Policy Change Rates Over Time of Each Agent in a Four-Agent Network Corresponding to Different Connectivity Structure: A) A Fully Connected Network; B) A Network with a Central Agent; C) A Ring Network; D) A Network with a Local Cluster. The Network Structures are Shown in Figure 3.8

Figure 4.3 shows the policy change percentage of each agent over training time in four different connectivity structures, in which we could see that the policy changes of all agents start from approximately 0.6 and decrease to around 0.05. The observation indicates that all agents could converge in a stable policy after a certain number of time steps, which is around 12000 for four-agent networks. Besides, the decreasing speed of the policy change rate for each agent in a community is similar despite the unique connectivity properties of each agent.

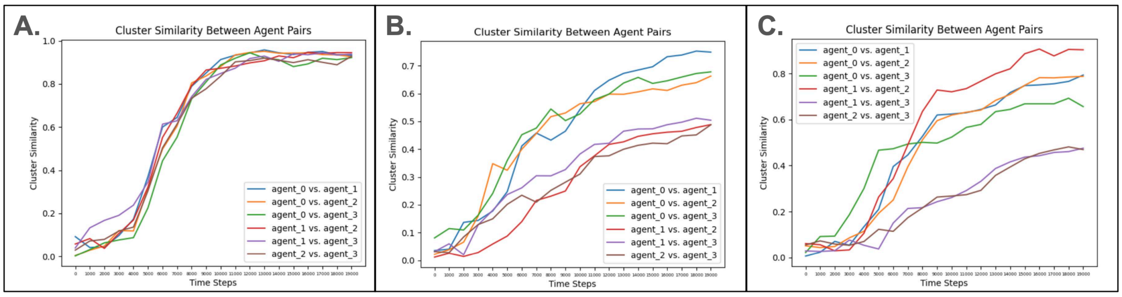

Figure 4.4: The Pair-Wise Policy Similarities Over Time in a Four-Agent Network Corresponding to Different Connectivity Structure: A) A Fully Connected Network; B) A Network with a Central Agent; C) A Network with a Local Cluster. The Network Structures are Shown in Figure 3.8

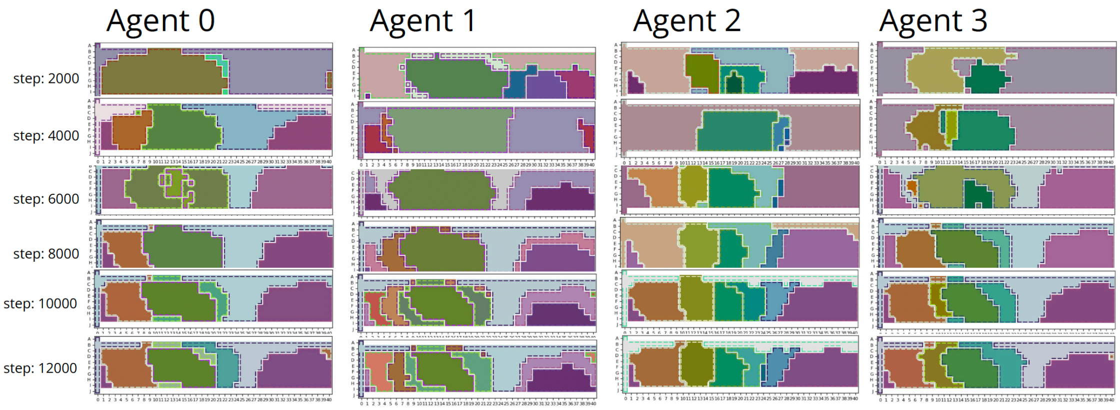

Figure 4.5: The Color Partitioning Patterns Changes of Four Agents through Communication in the Network with a Central Agent, as Shown in Figure 3.8 (C)

4.1.3 Exp 1-3: Communication within Community with Eight Agents

Then, we investigated how the connected relationship between two agents affects the pair-wise similarities, which could indicate the pair-wise communication efficiency. In the fully connected network, the pair-wise similarities are all at the same level, which satisfies our expectations since each agent connects with each other, while in the network with a central agent, agent 0, there are two categories of pair-wise similarities: those pairs with agent 0, in other words, the pair connects to each other directly, have higher similarities than those pairs in which no communicating edge connected each other. Notably, in the network with a local cluster composed of agent 0, agent 1 and agent 2, and a separate agent connecting with the local cluster through agent 0, the pair of agent 1 and agent 2 has the highest similarity, the possible explanation is that they connect with each other directly in a local cluster without communicating with another agent outside the cluster, while the similarities of pairs of agent 0 and agent 1 or agent 2 are slightly lower, since agent 0 needs to develop a language pattern that could enable the efficient communication both inside the cluster (with agent 1 and agent 2), and outside the cluster (with agent 3).

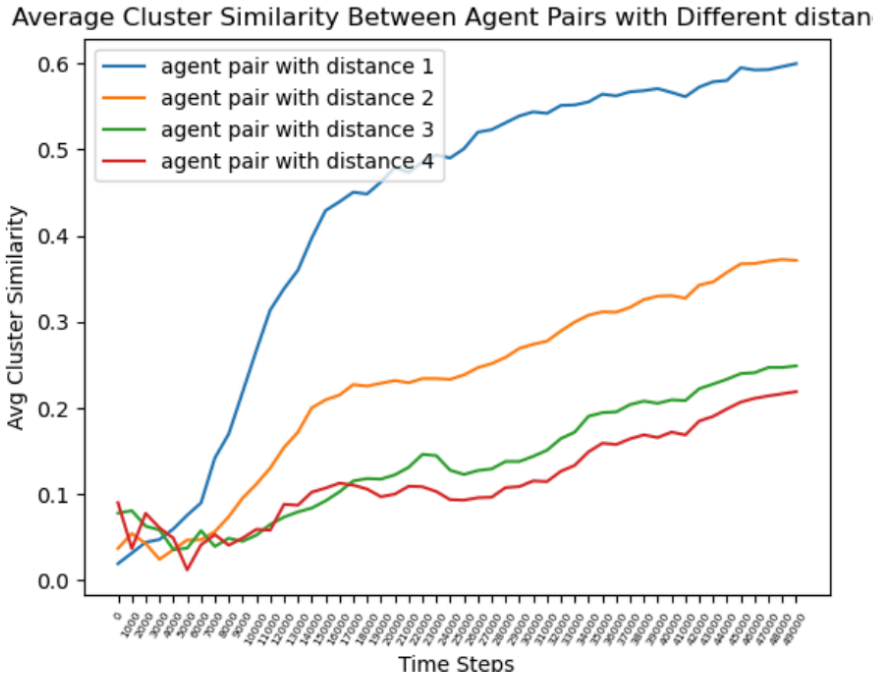

Figure 4.6: The Pair-Wise Policy Similarities Over Time in an Eight-Agent Ring Network.

Figure 4.5 shows how the four agents affect each other's policies. From Figure 4.1 and Figure 4.5, we could see that there are more similar partitioning patterns in the warm color region, which is the left and right area in the color map, while the cool color region in the central part of the color map has more distinctive patterns, which is similar with the observation from the color partitioning patterns in human language shown in Section 2.5: warm colors could be communicated more precisely than cool colors.

Inspired by the observation comparing the pairwise similarities in different connectivity relationships in Exp 1-2 in Section 4.1.2, we further investigated how the connected distance, which is the distance of the shortest path that connects two different agents, affects the policy similarity and communication efficiency. As shown in Figure 4.6, we could see the color cluster similarity between a pair of agents increase with the communicating distance between these two agents decreasing.

Furthermore, we investigated how the communication performance changes along with two conditional variables: the average connectivity degree within one community and the size of the given vocabulary. Figure 4.7 indicates the results varying the two conditional variables, from which we could observe that when the average connectivity degree increases, the reward, which indicates the communication accuracy, and the inner consistency increase in the same trend, which means if more agents in a community could communicate with each other directly, they tend to form a common language pattern, which helps them communicate more accurately and efficiently. Meanwhile, the number of unique used words decreases when the average connectivity degree increases (a word is regarded as a used word if it has the

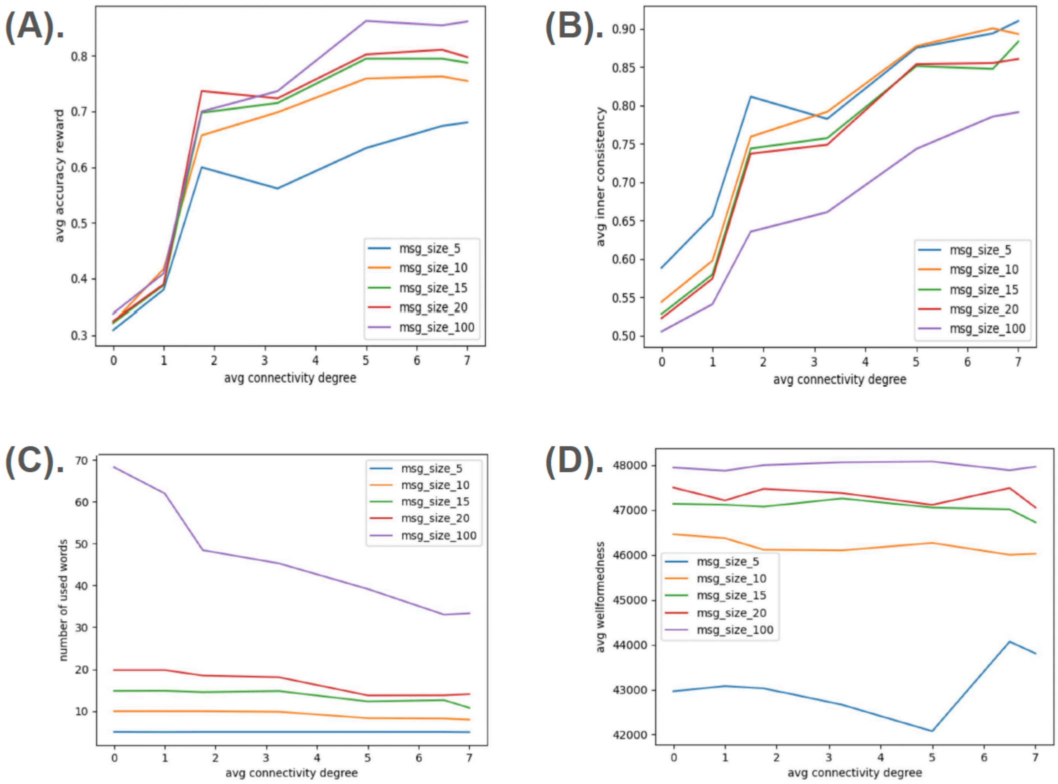

Figure 4.7: Communication Performance after Training and Converging Stable Communication Patterns Varies along with the Average Connectivity Degree within a Community. A) The Average Accuracy Reward, B) The Average Inner Consistency, C) The Average Number of Unique Words with Highest Probability for each Color, D) The Average Wellformedness.

highest probability conditional to a color), which illustrates the language system in a highly connected community tends to be less complex.

Figure 4.8 shows how the communication performance changes when the size of the given vocabulary increases. From Figure 4.8 (A), we observed that the communication accuracy is quite low, around 0.4, when we only provide the speaker with two optional words as the vocabulary. Because all 330 colors need to be represented by only two colors, the communication to distinguish colors is not efficient and not accurate. The communication accuracy increases with increasing the vocabulary size until it reaches 25, which means when language complexity reaches a threshold, increasing the language complexity wouldn't increase the communication accuracy, which is similar to the observation that in human emergent language and communication the trade-off between informativeness and simplicity tends to lie along the optimal or near-optimal, as described in Section 2.5. The same trend could also be seen in wellformed ness in Figure 4.8 (D), which indicates the rationality of color-partitioning patterns formed from communication in a community.

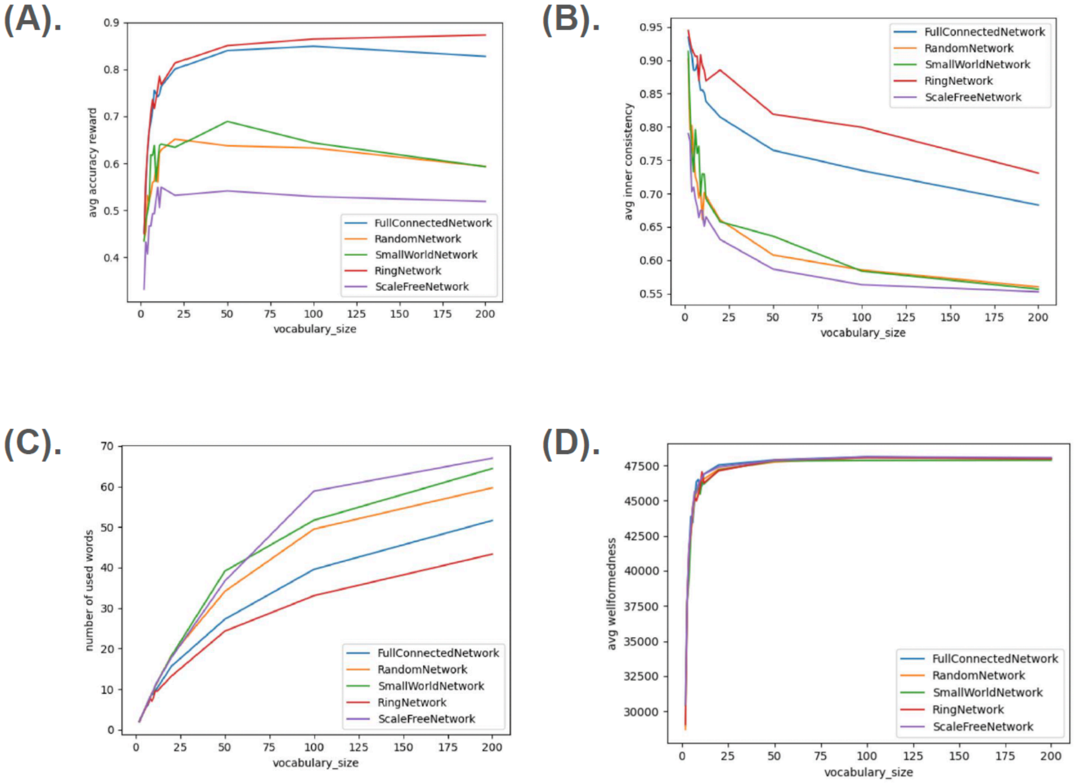

Figure 4.8: Communication Performance after Training and Converging Stable Communication Patterns Varies along with the Size of Given Vocabulary. A) The Average Accuracy Reward, B) The Average Inner Consistency, C) The Average Number of Unique Words with Highest Probability for each Color, D) The Average Wellformedness.

4.2 Communication between different communities

4.2.1 Exp 2-1: The impact of varying degrees of interaction between communities on the formation of new languages.

As we introduced in the previous experiment of Section 3.4.2.1, we adjusted the value of "Connect Num" to explore the performance of the two communities in language integration under different numbers of interactions. Specifically, we set the value range of the parameter "connect num" to [1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 56, 60, 64], where 64 is The maximum number of possible interactions based on the integration of two communities with a population size of 8 each.

In addition, as mentioned above, in the communication experiment of community integration, we used the speaker loss equation 3.2.2 and listener loss equation 3.2.2 and made differentiated settings for the β parameter. The setting of different β values

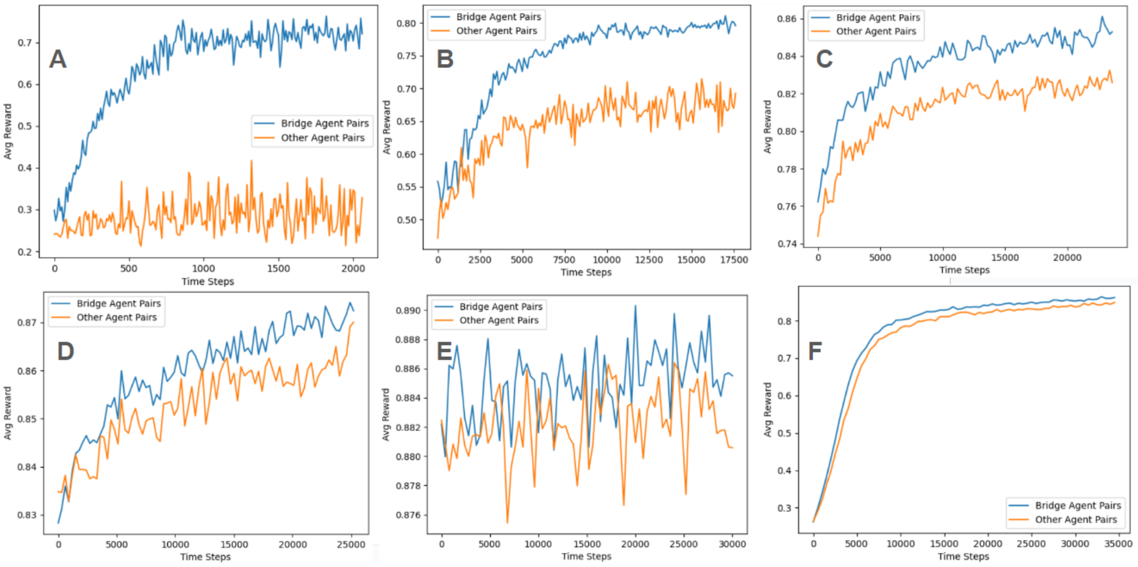

Figure 4.9: The average reward of "bridge-agent pairs communication" vs. "other-agent pairs communication". β is 0.01.The "connect num" values are 1 (fig A), 10 (fig B), 20 (fig C), 30 (fig D), 40 (fig E), 50 (fig F).

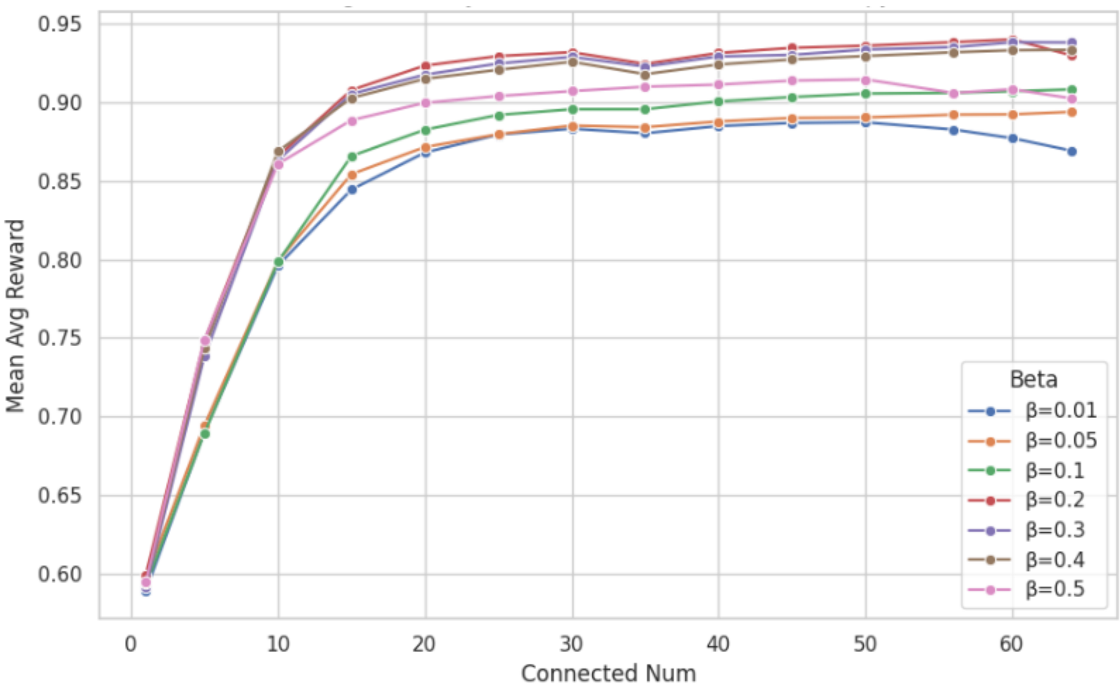

Figure 4.10: The average reward of all agents from the new community.

([0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5]) is designed to adjust the balance between exploring new language structures and maintaining existing language structures in the community during the language learning process of the two community-connected models. A higher β value can increase the entropy of the model output distribution, encourage the model to explore more diverse language expressions, and thereby promote effective communication between communities and the generation of new language forms. In contrast, a lower β value helps to strengthen the utilization of learned knowledge and maintains language consistency within the community, but may inhibit the exploration process of language integration.

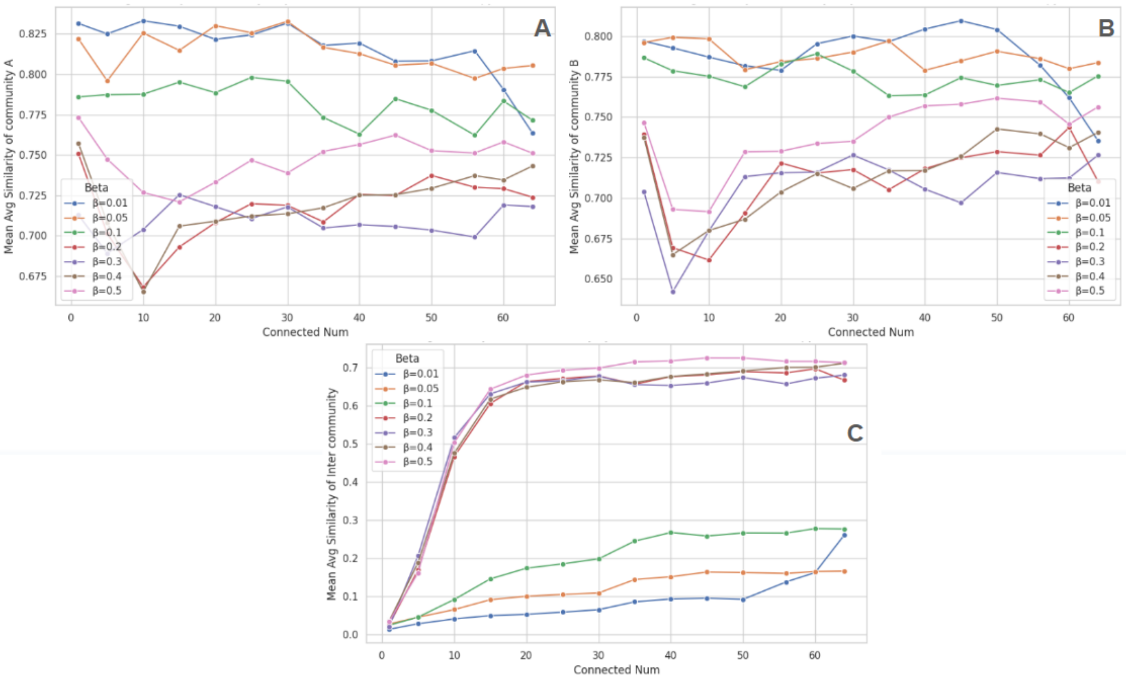

Figure 4.11: The Similarity of internal community A (fig A), internal community B (fig B) and interact community (fig C).

First, we set up two fixed and pre-trained communities: A and B. To evaluate the impact of each value of the "connect num" parameter in the list on the experimental results, we performed five experiments for each value. We calculated the average of the results to eliminate the influence of chance factors. In the experiment, we specifically recorded the communication effectiveness between "bridge-agent pairs" and "other-agent pairs."

As shown in Figure 4.9, the communication accuracy between "bridge agent communication" and "other agent communication" is shown when the "connect num" values are 1, 10, 20, 30, 40, and 50 and β is 0.01. According to the six subgraphs in the picture, we can observe that as "connect num" increases, the communication accuracy between "bridge agent pairs" and "other agent pairs" increases. Initially, the accuracy of "bridge agent pairs" is significantly higher than that of "other agent pairs." However, as the number of connections increases, this gap gradually narrows until the accuracy of the two converges. This shows that after reaching a sufficiently high number of exchanges (i.e., tending to converge), as "connect num" increases, the

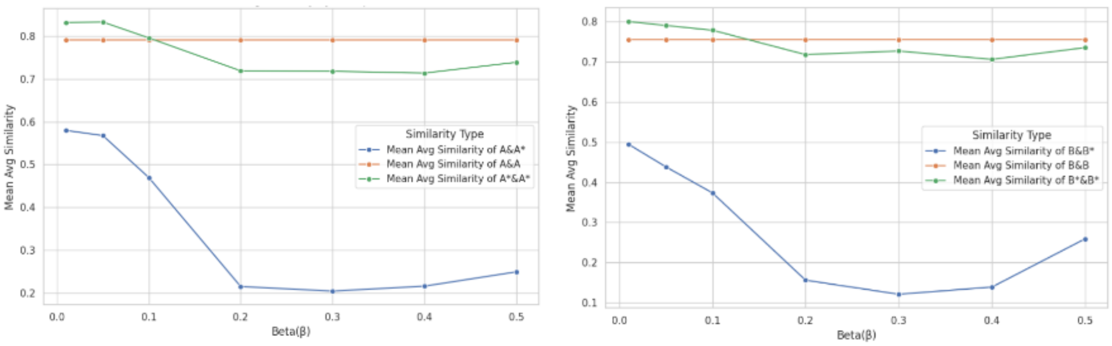

Figure 4.12: The average similarity within communities A, B, A* and B*.

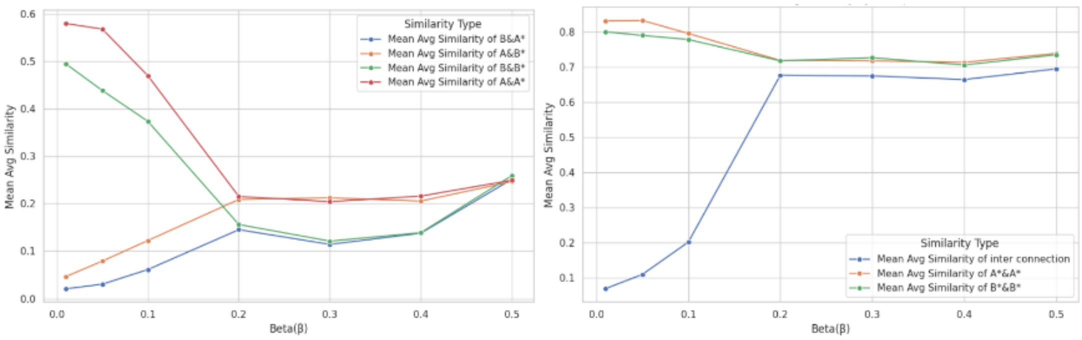

Figure 4.13: The average similarity between communities A, B, A* and B*.

language between the two communities begins to evolve from unification through connected agents to a comprehensive unification of the language of the entire community. This phenomenon shows that even if there are members in the community who are not directly connected and communicating, they can successfully understand each other's language.

After conducting a series of experiments on language communication with agents in a new community, we gained meaningful insights from the data. As shown in Figure 4.10, we observe that when "connect num" reaches 20 or higher, the average reward of all agents under different β settings shows higher communication accuracy. This suggests that a certain number of inter-connections can promote effective verbal communication, regardless of the specific value of β.

In Figure 4.11, we show in detail the average internal similarity between speaker and listener agents belonging to community A and community B after community fusion training and the average inter-community agent interaction similarity. The results show that low β values promote linguistic consistency within communities but have lower linguistic similarity between different communities. When the β value reaches 0.2 and above, the similarity between the two communities increases significantly, indicating that a new language form has been formed, promoting cross-community language integration.

In the context of community integration, we set the "number of connections" to 30.

We explored the internal and external relationships between communities A and B through Figure 4.12 and Figure 4.13 the language similarity changes between them. Experimental results show that when β is 0.2, the similarity within the community reaches the lowest and then gradually increases, which reflects that at moderate β values, new languages begin to form and are widely accepted by community members. At the same time, as the value of β increases, We found that the similarity between community A*&B* gets closer to A*&A* and B*&B*, and the similarity of A&A* gets closer to A&B* as the entropy increases. , This shows that increasing β does help bridge the language gaps between different communities and promote the formation of a unified language.

4.2.2 Exp 2-2: The impact of adding a new agent (newborn) between communities on the formation of new languages.

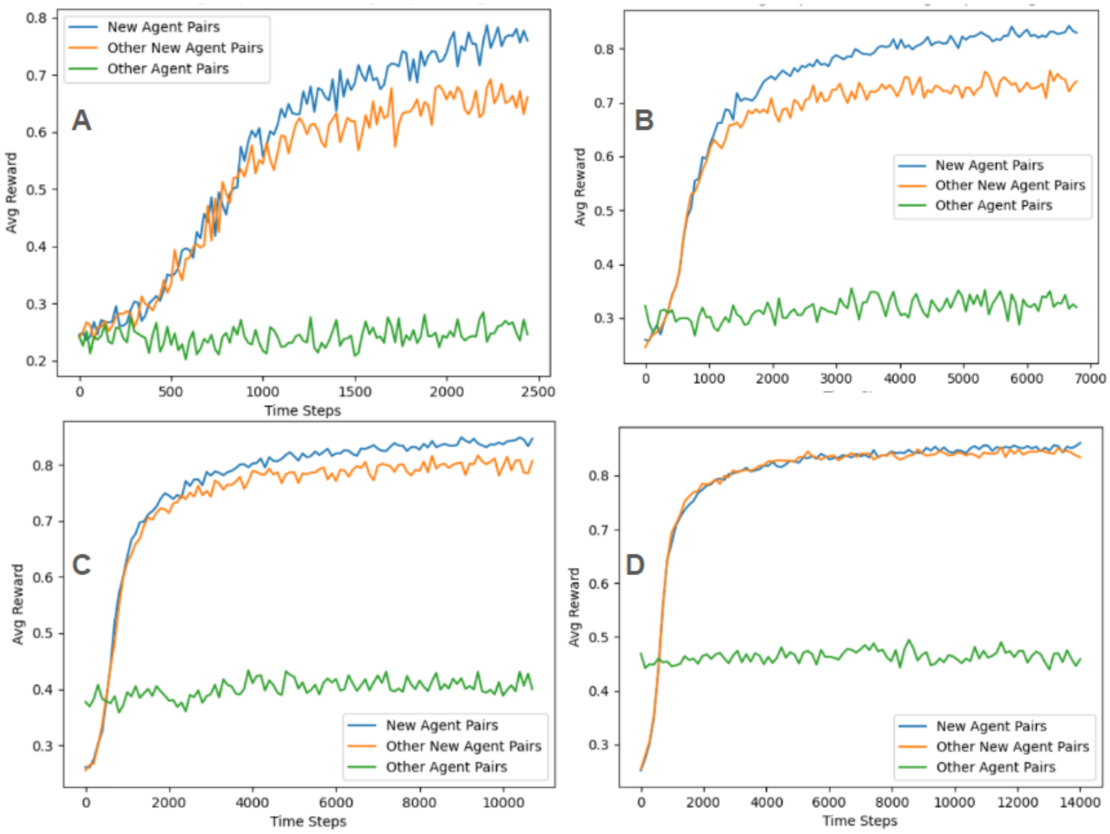

Figure 4.14: The average reward of "new agent pairs communication" vs. "other new agent pairs communication" vs. "other agent pairs". β is 0.01. The "new connect num" values are 2 (fig A), 6 (fig B), 10 (fig C) and 14 (fig D).

As described in our previous experiment 3.4.2.2, we introduced a new agent with the purpose of exploring the impact of "newborn" on community integration and the language formation process of the new agent in this integrated environment. Moreover, we adjusted the value of the "New Connect Num." Specifically, we set the parameter "new connect num" value range to [2, 4, 6, 8, 10, 12, 14, 16] because only

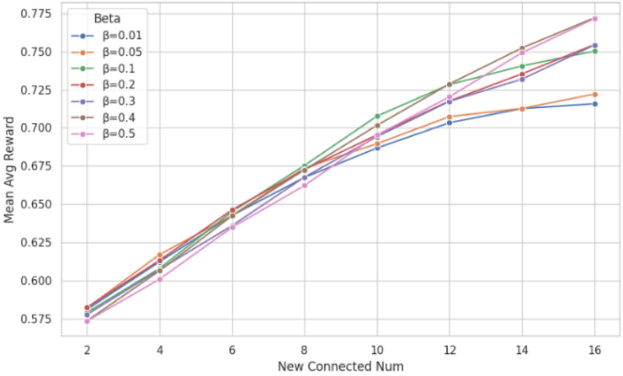

Figure 4.15: The average reward of all agents (including the new agent) from the new community.

up to 16 agents can connect to the new agent. In addition, we also adjust the β coefficient used to calculate the loss, which ranges from [0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5]

First, the experimental results (shown in Figure 4.14) reveal that the increase in the "number of new connections" significantly improves the communication accuracy of agent pairs directly connected to new agents. Although the accuracy of agent pairs that are not directly connected to the new agent also gradually improves, the agent pairs connected to the new agent always maintain higher communication efficiency. This suggests that new agents are vital in facilitating inter-community communication and language synchronization. In addition, it can be found that the communication accuracy of "other agent pairs" is much lower. The "other agent pairs" here are agent pairs composed of agents from two communities that are not connected to the new agent.

Furthermore, by observing the average reward changes in Figure 4.15, we find that as the number of connections increases, the communication accuracy of all types shows an upward trend but does not reach an equilibrium point. This means that although the number of connections has reached the maximum, there is still space for improvement in communication accuracy.

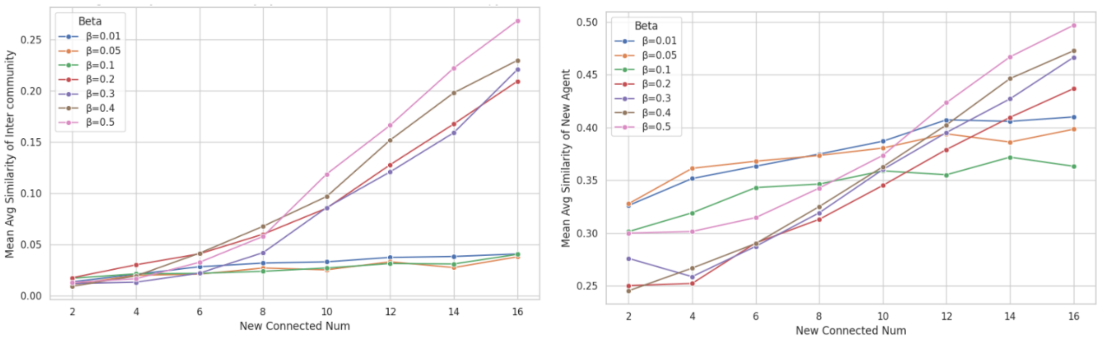

Figure 4.16: The Similarity of communities A and B (left fig), the similarity of the new agent and two communities (right fig).

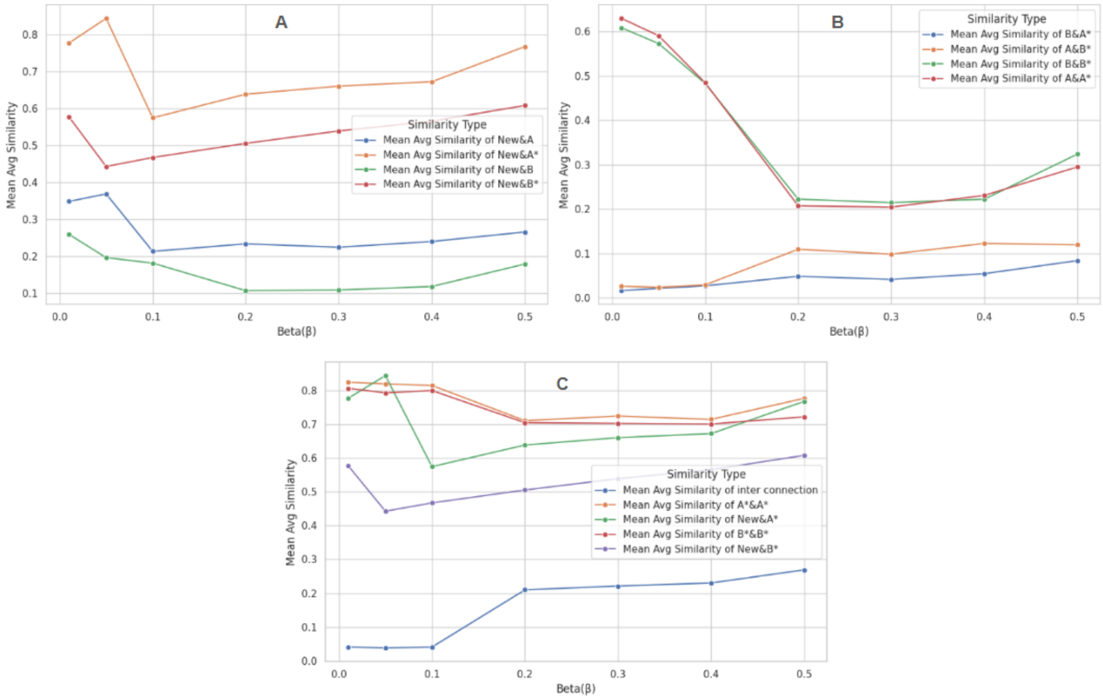

Figure 4.17: The average similarity within and between communities A, B, A*, B* and the new agent.

Regarding language similarity (as shown in Figure 4.16), when β is low, the similarity between communities A and B is still not high. However, the similarity between the two communities and the new agent is very high. When the "new connect num" reaches the maximum value of 16, the new agent's language shows the fusion of the language features of the two communities, and the average similarity of the two communities is close to 0.5, indicating that the new agent successfully integrates the two languages. , becoming a bridge for language integration.

Finally, Figure 4.17 has three subfigures, the first showing the high linguistic similarity between the new agent and the newly formed community. This shows that the new agent can effectively adapt and integrate into the new community environment, promoting language unification with new community members. The second subfigure shows the impact of β value on community language similarity, where 0.2 becomes an obvious watershed. When β reaches 0.2, the similarity between the newly generated community language and the language of the original community drops to a lower level and tends to be stable. The third graph is the similarity between communities A and B and the new agent, which shows that although the unique language characteristics between communities still exist, the language similarity between the new agent and each community is significantly higher than the similarity between communities. This shows that the new agent not only absorbs the language features of the two communities but also promotes language understanding and communication between the communities.

4.2.3 Exp 2-3: The effect of the ratio of the population sizes of two communities on the formation of new languages.

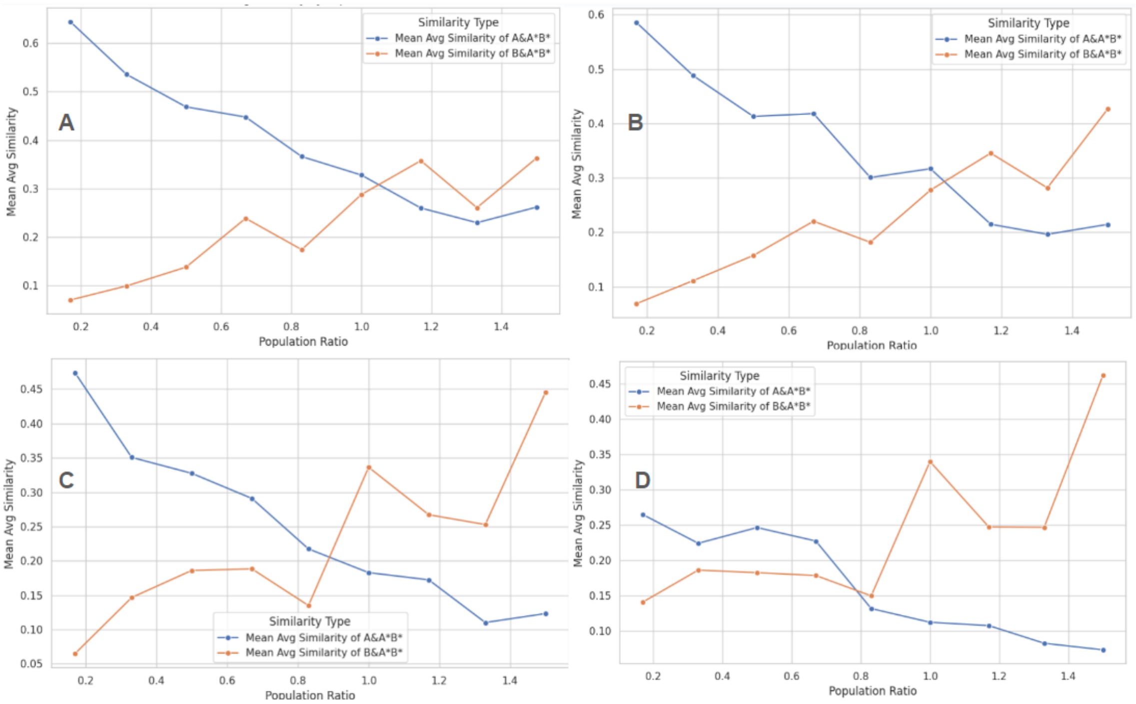

Figure 4.18: The similarity between the original community A community B and the new community after integration. The β values are 0.01 (fig A), 0.1 (fig B), 0.2 (fig C) and 0.5 (fig D).

As described in our previous experiment 3.4.2.3, we fixed the population size of community A to 12. Then we changed the population size of community B to [2, 4, 6, 8, 10, 12, 14, 16, 18]; the purpose is to explore the impact of different population ratios on community integration and whether adding new agents, in this case, will affect the language formation process. We adjusted the value of the parameter "Population Ratio." Moreover, experiments were set up in two situations: to add new agents and not to add new agents.

For the experiment of adding a new agent, we set the "new connect num" to the highest number of connections, which is the sum of the populations of the two communities. For experiments without adding a new agent, we set the "connect num" to half the maximum number of connections, which is 2/3 of all possible inter-agent pairs. In addition, we also adjust the β coefficient used to calculate the loss, which ranges from [0.01, 0.1, 0.2, 0.3, 0.5, 0.7].

According to the figure 4.18, we can find that in the direct connection experiment, as the population proportion increases (that is, the population size of community B increases), the language of the new community is more similar to that of community A to gradually become more similar to community B. We set different β values here. We found that no matter whether β is 0.01, 0.1, 0.2, or 0.5, the same trend will show;

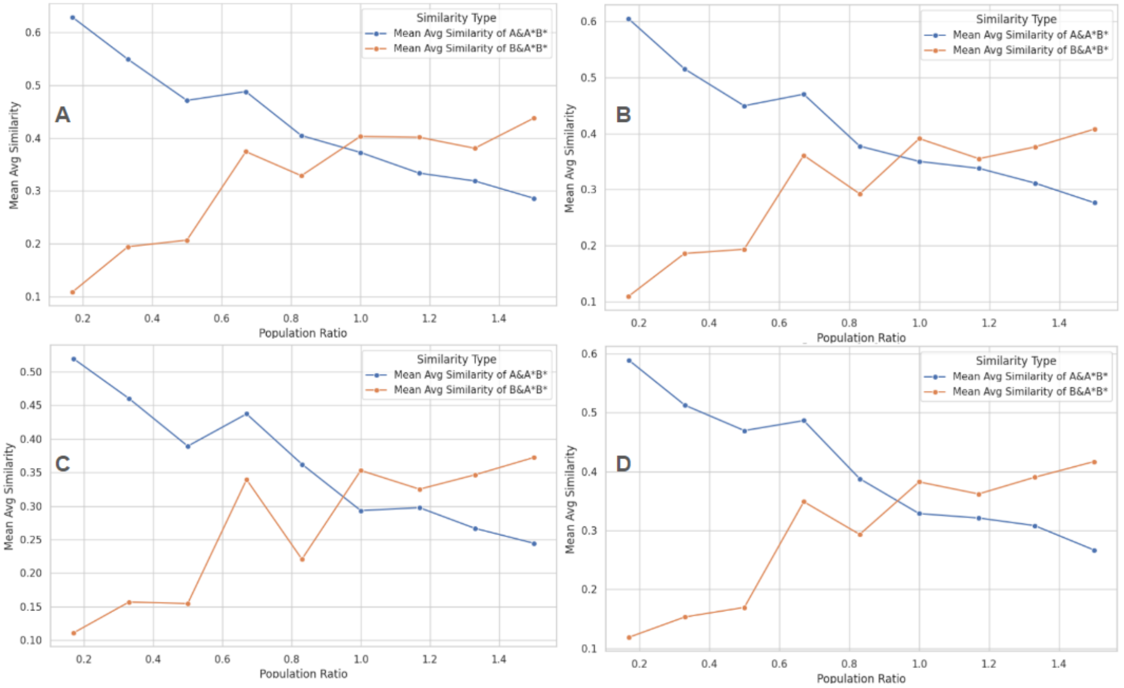

Figure 4.19: The similarity between the original community A community B and the new community after integration (with a new agent). The β values are 0.01 (fig A), 0.1 (fig B), 0.5 (fig C) and 0.7 (fig D).

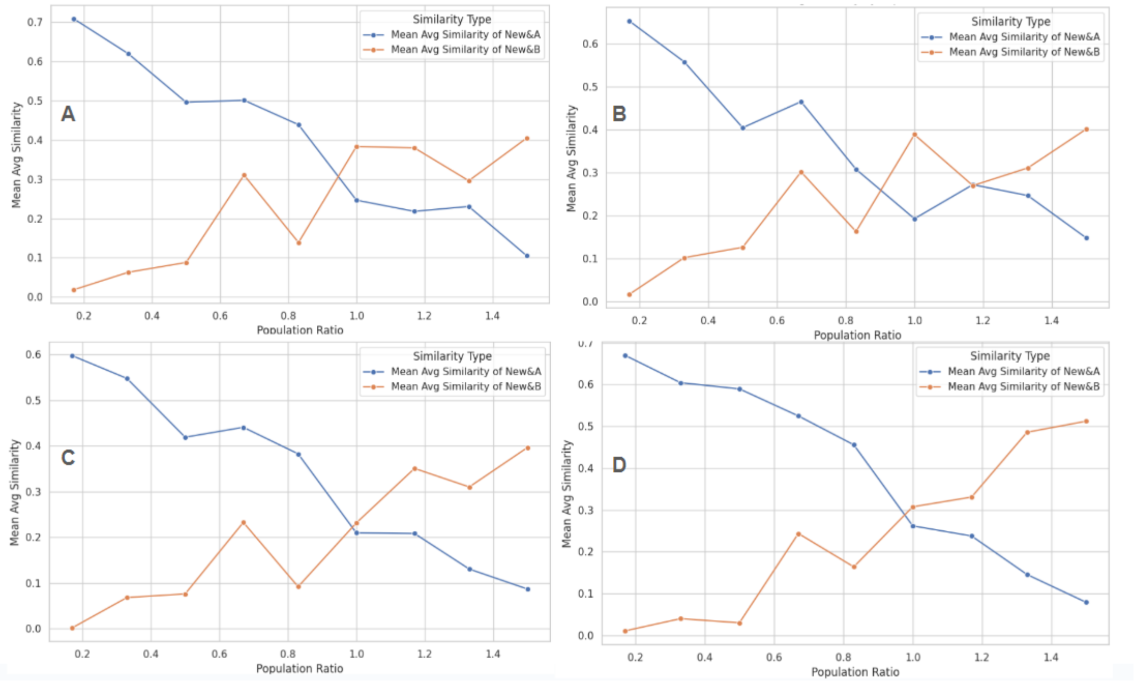

Figure 4.20: The similarity between the original community A community B and new agent. The β values are 0.01 (fig A), 0.1 (fig B), 0.5 (fig C) and 0.7 (fig D).

that is, when the population ratio is close to 1.0 (the population sizes of communities A and B are the same), the similarity will flip. This suggests that regardless of how β changes, the resulting language will be closer to a more populous community.

Subsequently, we conducted experiments in which a new agent was added to connect the two communities. According to Figure 4.19, we find that the changing trend of the curve is consistent with Figure 4.18. The language of the new community will still be closer to that of the more populous community. This shows that even if a new agent is added as a communication bridge rather than a direct connection, the generated language will still be closer to the more populated community, regardless of whether the value of β is low (0.01, 0.1) or high (0.5, 0.7).

Furthermore, Figure 4.20 shows the similarity between the new agents added in the second experiment and the original communities A and B. The picture shows that the curve trends in different subplots are very close as β changes. The language evolution pattern of the new agent is consistent with the evolution pattern of the two communities, and the color naming and classification of the new agent are also closer to the community with larger population size.

5

Discussion

5.1 Result conclusion

In this work, we successfully established a framework that allows multi-agent communication in the color domain and supports each agent in forming a unique language pattern that follows the human scenario. Based on previous work, we have made some improvements to this framework: each agent could play the role of speaker and listener, relying on the same policies. This framework could be extended and used in emergent language studies. At the same time, we gained some insights from multi-agent communication within one community and inter-community communication.

For the experiments involving multi-agent communication within one community, we visualized the language emergence process in different social networks. We discovered that when two agents connect closer, which means they can reach each other through a shorter path, they tend to form more similar language patterns. However, the difference in average similarity becomes insignificant when the distance between two agents is larger, as illustrated in Figure 4.6. This observation could relate to the phenomenon in human communication. Also, in the communication among artificial agents, we investigated the impacts of network properties (connectivity degree, size of vocabulary set) on the communication performance, such as accuracy, consistency, number of unique used words after convergence, and wellformedness of color clusters. Here are several valuable observations:

1. Agents in a closely connected network could form more consistent language patterns, as shown in Figure 4.7 (B).
2. The language system in a highly connected community tends to be less complex, as shown in Figure 4.7 (C).
3. Language complexity could increase the communication precision and the color partitioning quality until it reaches a threshold, as shown in Figure 4.8.

For the community communication experiments, we compared the communication accuracy based on different interaction connection numbers. As Figure 4.10 shows, we discovered that if sufficient inter-community connections are established, two communities with independent language systems can successfully communicate and undergo language fusion, ensuring both have ample opportunities for interaction.

However, as Figure 4.9 shows, when connections between communities are insufficient, language integration often remains limited to bridge agents serving as communication intermediaries and fails to extend broadly among community members. As shown in Figure 4.12 and Figure 4.13, successful language fusion in our experiments not only preserved some of the original ways both communities named and classified colors but also created new vocabularies for naming and classifying colors. This new vocabulary supplements and expands the original language classification and illustrates the mutual influence and innovation during language contact.

By introducing a new agent between the two communities, we observed that this agent successfully learned and merged both communities' color naming and classification structures while adding new vocabulary for naming colors (as shown in Figure 4.14). This phenomenon indicates that the new agent effectively integrates the language characteristics of both communities. However, Figure 4.16 shows that despite the new agent's integration of both languages, communication between the two original communities still needs to be improved, with a low success rate. As the number of connections to new agents increases, these agents further learn the languages of both communities (as shown in Figure 4.15). However, direct language integration between the two communities still requires more interaction and effort. Then, the comparison of Figure 4.17 and Figure 4.13 demonstrates that eliminating language barriers between two communities by adding new agents is more challenging than direct interaction between the communities.

As shown in Figure 4.18 and Figure 4.19, when studying communication between two communities with different population sizes, we found a clear trend: the color naming and categorization of new communities tended to be more similar to those of larger populations. This phenomenon occurs not only in direct inter-community communication but also when two communities are connected through new agents, reflecting the same pattern. In this setting, the new agent also showed higher similarity in color naming and classification with the more populous community.

Additionally, Figure 4.20 shows that when a new agent is introduced as a bridge connecting two communities, the similarity between the new agent and the original communities is usually slightly higher than the similarity between the new community formed directly without the new agent. This indicates that the new agent can learn and integrate the color naming and classification features of the two communities more effectively.

Here are several valuable observations:

1. Sufficient inter-community connections enable successful communication and language fusion between communities.
2. Insufficient connections limit language integration only to bridge agents, restricting broader community integration.
3. New agents can effectively integrate and expand the language systems of two communities.
4. Indirect integration through new agents requires more interaction and effort

than direct language integration between two communities to implement inter-community communication.

5. Larger communities significantly influence the language characteristics of new communities formed through integration.
6. New agents can more effectively learn and integrate the language features of both communities compared to direct community integration.

5.2 Limitation and future work

Despite the contributions we have made through this work, there are some limitations in our work and some extensive directions that remain to be explored further. We built a framework to simulate the communication between populations. However, we only tested this work on a limited size of population, which is different from the natural human community. This size limitation may result in research outcomes that do not accurately reflect the dynamics of large social populations. Therefore, future experiments should consider large-scale artificial agent setups, such as involving 50 or 100 agents. However, it is essential to note that every 80,000 exchanges take about 30 minutes in our current implementation. Improvements in experimental efficiency are necessary to implement large-scale agent experiments effectively.

Besides, in this work, we only look into communication between speaker-listener pair communication; there could be some work in implementing multiple-listener cooperation and analyzing how it affects language patterns. Meanwhile, we have visualized the policy changes of each agent in the training process and the color partitioning results after training. However, we still need to establish a comparison framework to compare them with the color partitioning patterns in human language.

In addition, there are indeed many possibilities for improvement in language experiments with community integration. For example, the integration of languages can be promoted by increasing the number of inter-community communication in the game, which may accelerate the creation of new languages. Similarly, adjusting the weight of intra-community and inter-community communication is also a helpful direction to explore. For example, according to research by Laura et al. [25], the importance of increased communication between communities is considered a critical factor in forming new languages.

In experiments that introduce new agents, we only add one agent; increasing the number of new agents may significantly impact the effectiveness of language fusion. By deploying multiple new agents, we can study whether multiple new agents can form a more effective language communication network and thus more effectively promote language integration among different communities. In addition, we can also consider reintroducing a new agent after the communication is stable and observe the iterative development of the language.

Different communication frequencies can also be considered when studying communities of varying population sizes to simulate real-world language exchange and integration processes. Additionally, one could explore methods to balance the linguistic influence between communities of different sizes, ensuring that communities with smaller populations have sufficient representation in the new language system.

6

Conclusion

This paper explores the development and evolution of emergent communication protocols between artificial agents using multi-agent reinforcement learning (MARL). By simulating various social network structures and leveraging the World Color Survey (WCS) color dataset, we can analyze how different connectivity patterns and community dynamics influence the formation of language-like systems.

An essential contribution of our work is introducing a shared neural network model for speakers and listeners, enhancing equivalence in language use between agents. Furthermore, we extend the traditional speaker-listener model to accommodate multiple agents to simulate more complex population game. Our results show that artificial agents can develop effective communication protocols through repeated interactions, resulting in emergent language patterns that reflect the structure of their social networks. Specifically, agents in more densely connected networks tend to develop more consistent and stable communication systems.

In addition, our experiments also highlight the impact of adding new agents and changing the degree of interaction between communities on language evolution, providing insights into the dynamics of language integration and the impact of population size scaling. Specifically, two communities can achieve successful communication when a certain number of interactions is reached. After introducing the new agent, the languages of both communities can be successfully merged into it. In addition, newly formed languages are closer to more populous communities.

However, our study has several limitations. First, our agent population was relatively small, which may not accurately reflect the complexity of larger social groups. To address this issue, future research should consider expanding the number of agents to simulate real-world scenarios better. Furthermore, our experiments show that running large-scale simulations is computationally expensive. Therefore, improving the efficiency of these experiments is crucial for practical large-scale applications.

Furthermore, the one-to-one communication model used in our experiments only partially replicates the complexity of real-world social interactions. Studying more complex language models that include one-to many or many-to-one communication and feedback loops can provide a more comprehensive understanding of emerging communication. Balancing the impact of language in communities of different population sizes through parameters such as communication frequencies is another area that could be explored in the future.

Bibliography

[1] W. Labov, "Some principles of linguistic methodology," *Language in society*, vol. 1, no. 1, pp. 97–120, 1972.

[2] M. Granovetter, "The strength of weak ties: A network theory revisited," *Sociological Theory*, vol. 1, pp. 201–233, 1983, ISSN: 07352751. [Online]. Available: `http://www.jstor.org/stable/202051` (visited on 05/20/2024).

[3] J. Milroy and L. Milroy, "Linguistic change, social network and speaker innovation1," *Journal of linguistics*, vol. 21, no. 2, pp. 339–384, 1985.

[4] L. Raviv, A. Meyer, and S. Lev-Ari, "The role of social network structure in the emergence of linguistic structure," *Cognitive Science*, vol. 44, no. 8, e12876, 2020. DOI: `https://doi.org/10.1111/cogs.12876`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12876`. [Online]. Available: `https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12876`.

[5] A. M. Ibrahim, K.-L. A. Yau, Y.-W. Chong, and C. Wu, "Applications of multi-agent deep reinforcement learning: Models and algorithms," *Applied Sciences*, vol. 11, no. 22, 2021, ISSN: 2076-3417. DOI: `10.3390/app112210870`. [Online]. Available: `https://www.mdpi.com/2076-3417/11/22/10870`.

[6] A. Lazaridou, A. Peysakhovich, and M. Baroni, "Multi-agent cooperation and the emergence of (natural) language," 2017. arXiv: `1612.07182 [cs.CL]`.

[7] D. Bouchacourt and M. Baroni, *How agents see things: On visual representations in an emergent language game*, 2018. arXiv: `1808.10696 [cs.CL]`.

[8] M. Kågebäck, E. Carlsson, D. Dubhashi, and A. Sayeed, *A reinforcement-learning approach to efficient communication*, Jul. 2020. DOI: `10.1371/journal.pone.0234894`. [Online]. Available: `https://doi.org/10.1371/journal.pone.0234894`.

[9] E. Carlsson, D. Dubhashi, and T. Regier, *Cultural evolution via iterated learning and communication explains efficient color naming systems*, 2024. arXiv: `2305.10154 [cs.CL]`.

[10] J. D. Thomas, R. Santos-Rodríguez, R. Piechocki, and M. Anca, "Multilingual agents through multi-headed neural networks," 2021. arXiv: `2111.11129 [cs.AI]`.

[11] C. Myers-Scotton, *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes.* Oxford University Press, Aug. 2002, ISBN: 9780198299530. DOI: `10.1093/acprof:oso/9780198299530.001.0001`. [Online]. Available: `https://doi.org/10.1093/acprof:oso/9780198299530.001.0001`.

[12] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction.* Cambridge, MA, USA: A Bradford Book, 2018, ISBN: 0262039249.

[13] R. J. Williams, "Simple statistical gradient-following algorithms for connectionist reinforcement learning," *Mach. Learn.*, vol. 8, no. 34, pp. 229–256, May 1992, ISSN: 0885-6125. DOI: `10.1007/BF00992696`. [Online]. Available: `https://doi.org/10.1007/BF00992696`.

[14] M. S. Zaïem and E. Bennequin, *Learning to communicate in multi-agent reinforcement learning : A review*, 2019. arXiv: `1911.05438 [cs.LG]`.

[15] R. S. Sutton, D. McAllester, S. Singh, and Y. Mansour, "Policy gradient methods for reinforcement learning with function approximation," in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K. Müller, Eds., vol. 12, MIT Press, 1999. [Online]. Available: `https://proceedings.neurips.cc/paper_files/paper/1999/file/464d828b85b0bed98e80ade0a5c43b0f-Paper.pdf`.

[16] M. Tan, "Multi-agent reinforcement learning: Independent vs. cooperative agents," in *Readings in Agents.* San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1997, pp. 487–494, ISBN: 1558604952.

[17] P. K. Sharma, R. Fernandez, E. Zaroukian, M. Dorothy, A. Basak, and D. E. Asher, "Survey of recent multi-agent reinforcement learning algorithms utilizing centralized training," in *Artificial intelligence and machine learning for multi-domain operations applications III*, SPIE, vol. 11746, 2021, pp. 665–676.

[18] L. Buoniu, R. Babuka, and B. De Schutter, "Multi-agent reinforcement learning: An overview," in *Innovations in Multi-Agent Systems and Applications - 1*, D. Srinivasan and L. C. Jain, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 183–221, ISBN: 978-3-642-14435-6. DOI: `10.1007/978-3-642-14435-6_7`. [Online]. Available: `https://doi.org/10.1007/978-3-642-14435-6_7`.

[19] C. Guestrin, M. Lagoudakis, and R. Parr, "Coordinated reinforcement learning," in *ICML*, Citeseer, vol. 2, 2002, pp. 227–234.

[20] A. W. Johnson and T. K. Earle, *The evolution of human societies: from foraging group to agrarian state.* Stanford University Press, 2000.

[21] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-worldnetworks," *nature*, vol. 393, no. 6684, pp. 440–442, 1998.

[22] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[23] P. Erds, A. Rényi, *et al.*, "On the evolution of random graphs," *Publ. math. inst. hung. acad. sci*, vol. 5, no. 1, pp. 17–60, 1960.

[24] P. Minnes, P. Carlson, M. Ann Mccoll, M. Lou Nolte, J. Johnston, and K. Buell, "Community integration: A useful construct, but what does it really mean?" *Brain Injury*, vol. 17, no. 2, pp. 149–159, 2003.

[25] L. Graesser, K. Cho, and D. Kiela, *Emergent linguistic phenomena in multi-agent communication games*, 2020. arXiv: `1901.08706 [cs.CL]`.

[26] F. Li and M. Bowling, "Ease-of-teaching and language structure from emergent communication," *Advances in neural information processing systems*, vol. 32, 2019.

[27] R. E. Wang, J. White, J. Mu, and N. D. Goodman, *Calibrate your listeners! robust communication-based training for pragmatic speakers*, 2021. arXiv: `2110.05422 [cs.CL]`.

[28] A. Grover, M. Al-Shedivat, J. K. Gupta, Y. Burda, and H. Edwards, *Learning policy representations in multiagent systems*, 2018. arXiv: `1806.06464 [cs.MA]`.

[29] I. Mordatch and P. Abbeel, *Emergence of grounded compositional language in multi-agent populations*, 2018. arXiv: `1703.04908 [cs.AI]`.

[30] D. Lewis, *Convention: A philosophical study.* John Wiley & Sons, 2008.

[31] N. Brandizzi, "Toward more human-like ai communication: A review of emergent communication research," *IEEE Access*, vol. 11, pp. 142 317–142 340, 2023, ISSN: 2169-3536. DOI: `10.1109/access.2023.3339656`. [Online]. Available: `http://dx.doi.org/10.1109/ACCESS.2023.3339656`.

[32] G. Dagan, D. Hupkes, and E. Bruni, "Co-evolution of language and agents in referential games," *arXiv preprint arXiv:2001.03361*, 2020.

[33] S. Havrylov and I. Titov, "Emergence of language with multi-agent games: Learning to communicate with sequences of symbols," *Advances in neural information processing systems*, vol. 30, 2017.

[34] L. Steels and M. Loetzsch, "The grounded naming game," in *Experiments in Cultural Language Evolution*, John Benjamins, 2012, pp. 41–59. [Online]. Available: `https://www.jbe-platform.com/content/books/9789027274953-ais.3.04ste`.

[35] Y. Gao, G. Chen, and R. H. M. Chan, "Naming game on networks: Let everyone be both speaker and hearer," *Scientific Reports*, vol. 4, no. 1, Aug. 2014, ISSN: 2045-2322. DOI: `10.1038/srep06149`. [Online]. Available: `http://dx.doi.org/10.1038/srep06149`.

[36] R. Chaabouni, E. Kharitonov, E. Dupoux, and M. Baroni, "Communicating artificial neural networks develop efficient color-naming systems," *Proceedings of the National Academy of Sciences*, vol. 118, no. 12, e2016569118, 2021. DOI: `10.1073/pnas.2016569118`. eprint: `https://www.pnas.org/doi/pdf/10.1073/pnas.2016569118`. [Online]. Available: `https://www.pnas.org/doi/abs/10.1073/pnas.2016569118`.

[37] D. T. Lindsey and A. M. Brown, "World color survey color naming reveals universal motifs and their within-language diversity," *Proceedings of the National Academy of Sciences*, vol. 106, no. 47, pp. 19 785–19 790, 2009. DOI: `10.1073/pnas.0910981106`. eprint: `https://www.pnas.org/doi/pdf/10.1073/pnas.0910981106`. [Online]. Available: `https://www.pnas.org/doi/abs/10.1073/pnas.0910981106`.

[38] N. Zaslavsky, C. Kemp, N. Tishby, and T. Regier, "Color naming reflects both perceptual structure and communicative need," *Topics in Cognitive Science*, vol. 11, Nov. 2018. DOI: `10.1111/tops.12395`.

[39] T. Regier, C. Kemp, and P. Kay, "Word meanings across languages support efficient communication," 2015. [Online]. Available: `https://api.semanticscholar.org/CorpusID:8291243`.

[40] ICSI, *World color survey*, `https://www.icsi.berkeley.edu/wcs/`, Partially Updated: 2017-02-27.

[41] X. Ohmer, M. Franke, and P. König, "Mutual exclusivity in pragmatic agents," *Cognitive Science*, vol. 46, no. 1, e13069, 2021.

[42] E. Carlsson and D. P. Dubhashi, "Pragmatic reasoning in structured signaling games," *ArXiv*, vol. abs/2305.10167, 2023. [Online]. Available: `https://api.semanticscholar.org/CorpusID:258740883`.

[43] R. Lowe, J. Foerster, Y.-L. Boureau, J. Pineau, and Y. Dauphin, *On the pitfalls of measuring emergent communication*, 2019. arXiv: `1903.05168 [cs.LG]`.

[44] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical association*, vol. 66, no. 336, pp. 846–850, 1971.

[45] T. Regier, P. Kay, and N. Khetarpal, "Color naming reflects optimal partitions of color space," *Proceedings of the National Academy of Sciences*, vol. 104, no. 4, pp. 1436–1441, 2007.