How Implicit Communication Emerges during Conversation Game

Extended Abstract

Bohao Wang
University of Tsukuba
Ibaraki, Japan
hiroforccat@gmail.com

Hirotaka Osawa
University of Tsukuba
Ibaraki, Japan
osawa@iit.tsukuba.ac.jp

Ken Satoh
National Institute of Informatics
Tokyo, Japan
ksatoh@nii.ac.jp

KEYWORDS

Multi-agent Simulation; Agent Communication; Conversation Games

ACM Reference Format:
Bohao Wang, Hirotaka Osawa, and Ken Satoh. 2018. How Implicit Communication Emerges during Conversation Game. In Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), Stockholm, Sweden, July 10–15, 2018, IFAAMAS, 3 pages.

1 INTRODUCTION

In this research, we used an evolutionary computation to investigate a linguistic phenomenon, implicit communication. Implicit communication is a phenomenon related to language evolution. This phenomenon performs as an ability to "read" an opponent's real intention from their superficial message [1]. Implicit communication generally means that talking has an implicit meaning which is shared by people in a subgroup [3]. This implicit communication is useful for achieving coordination in a subgroup, when the competitor is in a group. This type of implicit communication is expected to emerge and be changed during communication. For example, people in Japan's old city Kyoto use the phrase "recommend of simple dinner" as "recommendation to go home." This kind of inverted meaning of implicit communication shows who is an outsider of the group.

Implicit communication is expected to emerge according to the evolution of a language. For example, slang is a kind of implicit communication, that is generated in a subgroup of a community for communicating each other.

Implicit communication is also a smart strategy on conversation games. Lee et al. suggested that game players with implicit communication show better communication than players with a fixed mindset [4]. In this research, we study implicit communication with a conversation game, werewolf. Werewolf is a reinvention of a multi-player conversation game, Mafia [10][11]. In the werewolf game, players are distributed into two subgroups with hidden identity. One subgroup includes villagers. Villagers' purpose is to find out their enemies, werewolves in their opponent subgroup. In the villagers' group, a villager with special talent, the seer, has an ability to check one player's identity. Obviously, it is a common strategy that the seer shares the information about werewolves with other villagers. However, the seer must face two problems: Who is a villager? Did werewolves notice that I am the seer. The seer needs to find out villagers and persuade them to trust him as the seer, because werewolves usually pretend to be a seer for confusing villagers. On the other hand, if the seer exposed his identity to werewolves, he will be killed by werewolves in common strategy. That is not benefit to villagers. Thus, lie is a useful approach in werewolf game [7][8][9]. Returning to the implicit communication, villagers need to find out the outsiders in the group, which means implicit communication possibly emerges in villagers group.

The implicit communication in human-based werewolf game is researched on nonverbal information [5] and words used in discussion [6], how implicit communication emerges on agent communication is still a new study. In this article, we implement an evolutionary computation to investigate this study. Because the possibility of studying linguistic problems with agents has been shown, since Steel [2] supposed an approach, language game for agents, to study artificial language problems. For the werewolf game, the strong relationship between lie-based coordination and winning rate, and rule of hidden identity in the game indicated that implicit communication may be concealed in the players' irrational strategies and implicit lies. Therefore, we defined implicit communication as a player using irrational strategies by predicting that its cooperator is telling implicit lies. We classified our evolutionary computation with two conditions based on whether the werewolf agent is evolved. We hypothesized that when two cooperative agents were evolved, they will have an implicit communication in which the seer tells lies as a villager while the villager irrationally indicates that player who pretends to be a villager is the real seer. In the condition of three agents being evolved, implicit communication emerges in short periods and the used words are also evolved. On the other hand, the implicit interaction between participants reduces the need for an ambiguous argument [12] in conversation games. Therefore, we analyzed the results of the Werewolf evolutionary computation by labeling the agents' strategies as rational, irrational and ambiguous. Based on the analysis of the lie rate and the use rate of the classified agent's strategies, we explored the emergence of implicit communication

Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), B. Wang, H. Osawa, K. Satoh, July 10-15, 2018, Stockholm, Sweden.

from the view of a villager: a player without information on the competitors in its group.

2 THREE-PLAYER RULE OF WEREWOLF GAME

The simplest version of Werewolf comprises of a villager, seer, and werewolf. In this paper, we assumed the seer had identified the werewolf in advance. There was only a day phase and every player could speak only once. We modified the simplest version of Werewolf, to force a player to pretend to be a seer and provide real/fake information about who is to the werewolf. In the simplest version, there is a villager and a seer for the villager group and a werewolf for the werewolf group.

3 DESCRIPTION OF EVOLUTIONARY COMPUTATION

Definition of talks. In the regular version of Werewolf, a player can talk freely without restriction. Players can talk not about only things in the present, but also in the past or future. They can even discuss topics unrelated to the game. Theoretically, there are an infinite number of possible talks in a Werewolf game. Although researchers have made considerable progress in natural language processing, it is still difficult to analyze the complex Werewolf game with free talks. Therefore, we limited the players' available talks into 3 kinds of words,

1. "I am a villager."
2. "I am a seer. Werewolf is on my left side"
3. "I am a seer. Werewolf is on my right side"

Computation Process. Overall, the players decided their talks according to their hypothetical game status (genecode1) in Fig. 1. There were 2 kinds of game statuses as an input, and the output included 3 types of talks. The input includes two statuses: seer, werewolf and villager order in clockwise/counterclockwise. The output is 3 kinds of words as previously mentioned. Then, players updated their hypothetical game status based on the other two players' talks (genecode2). Since there were 3 types of talks for each player, there were 9 combinations of talks for the input of genecode2 and 2 kinds of game statuses, for the output. The output of genecode2 is the same form as the input of genecode1.

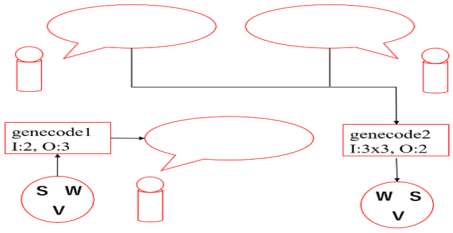

Figure 1: Process of the evolutionary computation.

4 CATEGORIES OF STRATEGY

The strategies of the players are 9 combinations of genecode2. We labeled the strategies as rational, irrational and ambiguous. A rational strategy is that a player recognizes the seer based on assuming the seer is telling the truth. If a player determined the seer based on assuming the seer was telling lies, it was an irrational strategy. When the player could not recognize the seer, we call it an ambiguous strategy.

5 RESULTS AND DISCUSSION

In the condition of two cooperative agents evolved, the situation of villager chooses irrational strategy and wins the game (purple part in the upper of Fig. 2) appears in 99.2% of generations. The seer kept tell lies as pretending to be a villager during all periods. In the condition of three agents evolved, implicit communication between seer and villager (purple part in the lower of Fig. 2) emerged in short periods (in 26.4% of generations). The seer's lies repeated in pretending to be a villager and indicating the villager is a werewolf.

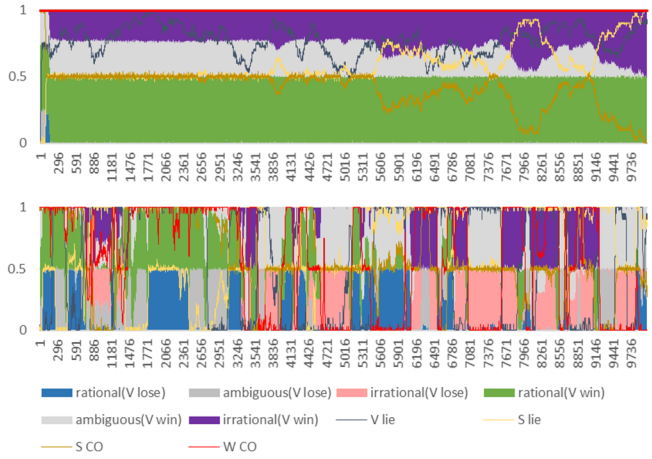

Figure 2: Lie rate, CO rate, and use rate of the villager's strategy for the condition of two cooperative agents evolved (upper) and three agents evolved (lower) (10000 generations, average in 100 computations)

6 CONCLUSIONS

We implemented evolutionary computations for werewolf game to explore the emergence of implicit communication. The results validated our hypothesis of that when two cooperative agents were evolved, they will have an implicit communication in which the seer tells lies as a villager while the villager irrationally indicates that player who pretends to be a villager is the real seer. In the condition of three agents being evolved, implicit communication emerges in short periods and the used words are also evolved.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP26118006, Artificial Intelligence Research Promotion Foundation, and Hayao Nakayama Foundation for Science & Technology and Culture.

REFERENCES

[1] Castelfranchi, C., Pezzulo, G., & Tummolini, L. (2012). Behavioral Implicit Communication (BIC): Communicating with Smart Environments via our Practical Behavior and Its Traces. In *Innovative Applications of Ambient Intelligence: Advances in Smart Systems* (pp. 1–12). doi:10.4018/978-1-4666-0038-6.ch001

[2] Steels, L. (2001). Language games for autonomous robots. IEEE Intelligent systems, 16:16–22.

[3] Beuls, K., & Steels, L. (2013). Agent-Based Models of Strategies for the Emergence and Evolution of Grammatical Agreement. *PLoS ONE*, *8*(3). doi:10.1371/journal.pone.0058960

[4] Lee, Y.-H., Heeter, C., Magerko, B., & Medler, B. (2012). Gaming Mindsets: Implicit Theories in Serious Game Learning. *Cyberpsychology, Behavior, and Social Networking*, *15*(4), 190–194. doi:10.1089/cyber.2011.0328

[5] Kobayashi, Y., Osawa, H., Inaba, M., Shinoda, K., Toriumi, F., & Katagami. F., (2014). Development of werewolf match system for human players mediated with lifelike agents. In Proceedings of the second international conference on Human-agent interaction - HAI '14, 205–207. https://doi.org/10.1145/2658861.2658923

[6] Zhou, L., & Sung, Y., 2008. Cues to Deception in Online Chinese Groups. In Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008), 146–146. https://doi.org/10.1109/HICSS.2008.109

[7] Bi, X., & Tanaka, T. (2016). Human-side strategies in the werewolf game against the stealth werewolf strategy. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 10068 LNCS, pp. 93–102). doi:10.1007/978-3-319-50935-8_9

[8] Katagami, D., Takaku, S., Inaba, M., Osawa, H., Shinoda, K., Nishino, J., & Toriumi, F. (2014). Investigation of the effects of nonverbal information on werewolf. In *IEEE International Conference on Fuzzy Systems* (pp. 982–987). doi:10.1109/FUZZ-IEEE.2014.6891847

[9] Vázquez, M., Carter, E. J., Vaz, J. A., Forlizzi, J., Steinfeld, A., & Hudson, S. E. (2015). Social Group Interactions in a Role-Playing Game. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*, 9–10. doi:10.1145/2701973.2702039

[10] Hung, H., & Chittaranjan, G. (2010). The Idiap Wolf Corpus: Exploring Group Behaviour in a Competitive Role-playing Game. In *Proceedings of the 18th ACM International Conference on Multimedia* (pp. 879–882). New York, NY, USA: ACM. doi:10.1145/1873951.1874102

[11] Kobayashi, Y., Osawa, H., Inaba, M., Shinoda, K., Toriumi, F., & Katagami, D. (2014). Development of Werewolf Match System for Human Players Mediated. *Proceedings of the second international conference on Human-agent interaction*, 205–207. doi:10.1145/2658861.2658923

[12] Bergström, K. (2010). The Implicit Rules of Board Games – on the particulars of the lusory agreement. *Proceedings of the 14th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '10*, 86–93. doi:10.1145/1930488.1930506