

# Werewolf Game Modeling Using Action Probabilities Based on Play Log Analysis

Yuya Hirata<sup>1</sup>, Michimasa Inaba<sup>1</sup>(✉), Kenichi Takahashi<sup>1</sup>, Fujio Toriumi<sup>2</sup>,  
Hirotaka Osawa<sup>3</sup>, Daisuke Katagami<sup>4</sup>, and Kousuke Shinoda<sup>5</sup>

<sup>1</sup> Hiroshima City University, 3-4-1 Ozukahigashi, Asaminami-ku, Hiroshima, Japan  
hirata.y@cm.info.hiroshima-cu.ac.jp, {inaba,takahashi}@hiroshima-cu.ac.jp

<sup>2</sup> The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

<sup>3</sup> The University of Tsukuba, 1-1-1 Tenno-dai, Tsukubashi, Ibaraki, Japan

<sup>4</sup> Tokyo Polytechnic University, 1583 Iiyama, Atsugi-shi, Kanagawa, Japan

<sup>5</sup> The University of Electro-Communications, 1-5-1 Chifugaoka,  
Chofu, Tokyo 158-8585, Japan

**Abstract.** In this study, we construct a non-human agent that can play the werewolf game (i.e., AI wolf) with aims of creating more advanced intelligence and acquire more advanced communication skills for AI-based systems. We therefore constructed a behavioral model using information regarding human players and the decisions made by such players; all such information was obtained from play logs of the werewolf game. To confirm our model, we conducted simulation experiments of the werewolf game using an agent based on our proposed behavioral model, as well as a random agent for comparison. Consequently, we obtained an 81.55% coincidence ratio of agent behavior versus human behavior.

**Keywords:** Werewolf game · Communication game · Player modeling · Multi-player

## 1 Introduction

In the Werewolf Game (also called Mafia), werewolves appear in a village in the form of humans during the day, attacking villagers one-by-one every night. The villagers decide that they must execute those who are suspected of being werewolves, but first must determine via discussions which villagers are actually werewolves. Since villagers are not given any information about others, information gleaned via discussion provides the only clue. Underlying the game is how human players see through the identities of the werewolf players and how werewolf players deceive the villagers, thus hiding their identities by providing limited information.

Studies in game informatics that began with chess have expanded in recent years, including games such as Go, Curling [1], and real-time simulation (RTS) [2]. Unlike in the case of them, the werewolf game differs in a way such that communication skills determine victory or defeat. The game requires the

use of our advanced intellectual ability, including the ability to understand the intention of others only from conversation, to therefore deduce the background of an individual and determine his or her willingness to cooperate or be persuaded. Thus, the werewolf game includes numerous communication-related problems that are significant barriers that prohibit artificial intelligence from penetrating the future society.

We, therefore, study agents that can play the werewolf game (i.e., AI wolf) with aims of creating more advanced intelligence and acquiring more advanced communication skills for AI-based systems. Construction of an AI wolf has identified numerous problems including persuading others to obtain trust, deducing an opponent model from information gleaned only from conversations, understanding and expressing non-verbal information, co-operating with other players, and applying natural language processing. We designed a protocol to discuss by AI wolves [3], released a construction kit of an AI wolf using the given protocol [4], and constructed the system that can compete with each other by anthropomorphic agent on the werewolf game. These initial studies aimed at solving the AI-specific problem of acquiring advanced communication skills by creating an environment where AI wolves can play the werewolf game and gathering collective intelligence via competitions.

To this end, in August 2015, we held the first AI wolf competition in The Computer Entertainment Developers Conference (CEDEC2015), which is the largest Japanese technical conference for game developers and engineers in Yokohama. More than 50 teams participated in the competition. We also analyzed human gameplay to obtain knowledge for realizing the AI wolf. For example, we investigated the effects of non-verbal information in the werewolf game [5], revealing gestures that impact victory or defeat by analyzing videos wherein human gestures were annotated in conjunction with the werewolf game. All these studies focused on constructing a strong agent for an AI wolf; such an agent requires not only strength but also the ability to behave much like humans.

Herein, we endeavor to realize an agent that can behave like humans by obtaining behavioral information from play logs of games played between humans, thus constructing a sound behavioral model. We used game situations obtained from play logs and action information of players in game situations to construct our behavioral model. Specific game situations and attributes, for example, the number of living villagers, number of players which have special ability, etc., can easily be obtained using the play logs; however, we could not obtain some information without analyzing utterance logs via natural language processing. We, therefore, obtained two types of information: (1) coming-out (CO) information: describing who, when, and what role is expressed, e.g., “Player A, on the first day, expressed him/herself as a seer”; and (2) decision information: describing who, when, and who was identified as either a villager or a werewolf, e.g., “Player B, on the third day, detected player C as a werewolf.”

## 2 What is the Werewolf Game?

### 2.1 Gameplay

In the game, all players are randomly allocated to roles, as summarized in Table 1. Players are divided into two teams: humans and werewolves. To win as a human, the goal is to kill all werewolves, whereas to win as a werewolf, the goal is to kill humans to the number of werewolves or fewer. Players do not know what the other players' roles are, as the assigned roles are hidden. A basic course of action of a human player is to find werewolves via conversation. Conversely, werewolf players know who all of the werewolves are. A basic course of action for a werewolf player is to engage in a variety of cooperative maneuvering without other humans learning of their role.

In the werewolf game, there are two phases: day and night. In the day phase, all players discuss who the werewolves are. Players who have special abilities, as described below, lead discussions to gain an advantage for their team by using the information gained via their abilities. After a certain period of time, players identify and execute one player suspected as a werewolf; this player is selected via a collective vote. The executed player cannot play the game from then on. In the night phase, werewolf players attack human players. Attacked players are also eliminated from the game. Moreover, players with special abilities can use their abilities in the night phase. The day and night phases repeat until one group meets the conditions for winning. A crucial aspect of the werewolf game is for human players to detect the lies put forth by werewolf players. Persuading other players by using information given by their special abilities is also important. For werewolf players, the crucial aspect is to manipulate discussions to their advantage. Occasionally, werewolf players must impersonate a role.

### 2.2 Roles

There are many variations in the werewolf game, often including roles with special abilities. Herein, we have adopted the orthodox roles especially in Japan. Table 1 shows these roles and the ability of each role. Seers have the ability to identify werewolves in the night phase, thus comprising the most important role in the werewolf game. Counter to this, werewolf players often impersonate the role of seers to disrupt and confuse the discussion; it is not uncommon for there to be three players impersonating the role of seers in a given game. Herein, we therefore obtained CO information and decision information from utterance logs.

## 3 Werewolf BBS

We used data obtained from the Werewolf BBS<sup>1</sup>, wherein users can play the game online. This site also provides discussion forums; overall, we could obtain utterance logs for all players, use history of special abilities, the role of each

<sup>1</sup> <http://nininix.x0.com/wolf0/>.

**Table 1.** Roles in the Werewolf Game.

Roles	Special abilities
Villager	Nothing
Werewolf	In the night phase, werewolf players attack one human player; werewolf players can attack one player in the whole of werewolf side in every night phase; werewolf players can secretly talk among themselves
Seer	In the night phase, a seer can select one player and know whether the selected player is a werewolf or not
Bodyguard	In the night phase, a bodyguard can select one player and protect the selected player from attack by a werewolf
Medium	A medium can know whether a player who is executed via a vote is a werewolf or not
Possessed	A possessed is a human, but his or her objective is the same as that of the werewolf players; a possessed is judged as human by both a seer and a medium (and is counted as a human player when victory or defeat judgments are made)

player, information of who is dead or alive at each date, and information as to the cause of death (i.e., by execution or attack); however, expressing one’s own role (i.e., CO) and speaking about other players is not included. Obtaining information as to whether a player is a werewolf is a special ability, but the timing in telling other players depends on the player. Moreover, there are players who impersonate other roles. Accordingly, we analyzed utterance logs using natural language processing to obtain such information. Note that it is difficult to obtain this information because utterance logs are written in various colloquial styles. Therefore, we created numerous regular expressions to obtain CO and decision information from utterance logs to cover many such variations.

## 4 Acquisition of CO and Decision Information

### 4.1 Using Regular Expressions

We obtained CO and decision information by using regular expressions. To construct regular expressions, considering the style and expressions used (on the utterance is included CO and decision information) are efficient; however, such an approach required tremendous time and cost to obtain utterances concerning CO and decision information by hand from all utterances in the Werewolf BBS.

Therefore, we obtained information regarding “when did seer (or medium) use special ability”, “which side did seer (or medium) decide”, and “who did seer (or medium) decide” by checking the use history of special abilities employed by players in Werewolf BBS. Next, we obtained the utterance that includes the target player’s name and result from utterances spoken by the seer and medium players on the day the special ability was used. These utterances included CO

**Table 2.** Example utterances.

Utterance	Type	Regular expression
I am seer	CO	⟨⟨USER⟩⟩ is ⟨⟨ROLE⟩⟩
I confirm CO of seer of Albin	CO	I confirm CO of ⟨⟨ROLE⟩⟩ of ⟨⟨USER⟩⟩
Peter is werewolf	decision	⟨⟨USER⟩⟩ is ⟨⟨DECISION⟩⟩

and decision information; we could obtain CO information because a decision was often reported at the same time as that of the CO. We, therefore, constructed regular expressions to obtain CO and decision information by using these utterances for reference.

As we analyzed the play log, for example, when player A is identified as a seer, other players sometimes also stated that “Player A is a seer.” Using such utterances, we could also obtain CO information from players without special abilities. Therefore, we used other players’ utterances on the same day to construct regular expressions. As such, we constructed 477 regular expressions and obtained CO and decision information via those regular expressions.

An example of utterances and regular expressions is shown in Table 2. In the table, ⟨⟨USER⟩⟩ accepts first-person pronouns and a player’s name, ⟨⟨ROLE⟩⟩ accepts names of roles, and ⟨⟨DECISION⟩⟩ accepts words that represent one side or the other, i.e., “werewolf” or “human” or the like.

## 4.2 Performance Evaluation of Regular Expressions

We conducted an experiment to evaluate the performance of our CO and decision information acquisition method via regular expressions. In this experiment, we randomly selected 50 games for CO information and 10 for decision information. We also evaluated our method using CO and decision information acquired by hand.

We evaluated the performance by measuring precision, recall, and the f-measure. If CO information obtained via the regular expressions completely matched what we acquired by hand, we noted the obtained CO information as correct. Similarly, if decision information obtained via the regular expressions completely matched what we acquired by hand, we noted the obtained decision information as correct.

Consequently, there were 193 utterances containing CO information in 50 games. From our regular expressions, 193 were matched and 190 were correct. Furthermore, there were 156 utterances containing decision information in 10 games. By our regular expression, 137 were matched and 114 were correct. Results are shown in Table 3, from which we observe that CO information yielded a very high precision and recall. Results regarding decision information were worse than those of CO information. In the case of decision information, depending on whether the speaker had a special ability, the meaning of utterances (e.g., “Player A is a werewolf”) was changed to a decision based on either

**Table 3.** Precision and recall rates.

	Precision	Recall	F-measure
CO	98.4(190/193)	98.4(190/193)	0.98
Decision	83.2(114/137)	73.1(114/156)	0.78

their ability or just speculation. Thus, obtaining decision information was more difficult than obtaining CO information, as is evident in our experimental results.

## 5 Behavioral Model Based on Action Selection Probabilities

Here, we describe a method used to construct our behavioral model using data obtained from the Werewolf BBS. Our proposed model is targeted only at the behaviors and utterances shown in Table 4.

We define the probability that a player performs action  $a$  ( $a \in A$ ) in situation  $s$  ( $s \in S$ ) by the following equation:

$$p(a|s) = \frac{n_{s,a}}{\sum_{a \in A} n_{s,a}} \quad (1)$$

Here,  $p(a|s)$  represents the action selection probability and  $n_{s,a}$  is the number of times a player has performed action  $a$  in situation  $s$  in the given play logs. Situation  $s$  is defined based on the basis of a decision result reported by a player who is identified as a seer or medium and number of player which expressed him/herself as a seer or medium. For cases 1 and 2 of Table 4,  $A = \{\text{CO}, \text{not CO}\}$ . For cases 3, 4, 6, 7, and 8 from Table 4,  $A = \{p_1, p_2, \dots, p_k\}$ , assuming  $p_i$  is a player and  $k$  is the number of players. Here,  $p_i$  is defined by the number of CO players, the CO type (i.e., seer or medium), and the reported results of their decision. For case 5 from Table 4,  $A = \{\text{human side player}, \text{werewolf side player}\}$ .

To clarify our model, we describe a specific example wherein we focus on the selection of an attack target of a werewolf. Given that there are 10 players and a situation  $s$ , as described in Table 5, the game situation consists of two seers, a werewolf or a possessed expressing him/herself as a seer; furthermore, there are two players inspected as humans by those players who expressed themselves as seers. Other players did not express themselves and were not inspected by any special abilities. Here, executable action  $a$  (i.e., the player that can be an attack target) is shown in Table 6, but the werewolf is not included in action  $a$  because a werewolf cannot attack another werewolf. We assume that the numbers of occurrences of each action that the player took in the same situation in the play log were  $n_{s,p_1} = 854$ ,  $n_{s,p_2} = 3077$ , and  $n_{s,p_3} = 1320$ . We then obtained  $p(p_1|s) = 854/(854 + 3077 + 1320) = 854/5251 = 0.163$ ,  $p(p_2|s) = 3077/5251 = 0.586$ ,  $p(p_3|s) = 1320/5251 = 0.251$  by Eq. (1). Accordingly, given the situation of Table 5, the probability that a werewolf attacks a player who expressed him/herself as a seer is 16.3%; as for the player who was

**Table 4.** Modeled actions.

Num	Action
1	Whether a seer (or medium) expresses his/her role
2	Whether a werewolf (or possessed) expresses his/her fake role
3	Selection of divination target by a seer
4	Selection of a target possessing a special ability by a werewolf (or a possessed) impersonating a seer
5	Selection of a decision result of a special ability by a werewolf (or a possessed) impersonating a role
6	Selection of a vote target
7	Selection of an attack target by a werewolf
8	Selection of a protected target by a bodyguard

**Table 5.** Examples of situation  $s$ .

Factor	number of players
The number of players identified as a seer	2
The number of players inspected twice as human players by a seer	2
The number of players who have not been identified as any specific type and have not been inspected by a seer	6

inspected two times as a human player by a seer, this probability is 58.6%; and for the player who did not express him/herself and was not inspected by a seer, the probability is 25.1%.

As to why such probabilities occurred, the player who was inspected two times as a human player by a seer is trusted by human players because the player is very likely to be a human player. If there is a player who can be trusted, human players can advantageously discuss because by leading discussion by the player, werewolf players face increased difficulty in disrupting such discussions. Thus, the werewolf players prefer to attack players inspected two times as human players by a seer.

**Table 6.** Examples of action  $A$ .

$A$	Target player
$p_1$	The player who has been identified as a seer
$p_2$	The player who was inspected twice as human by a seer
$p_3$	The player who was not identified as any specific type and was not inspected by a seer

## 6 Degree of Coincidence Between Agents and the Play Logs

### 6.1 Outline

In this section, we investigate the degree of coincidence between agents using action selection probabilities and human behaviors from play logs. For comparison, we also created a random agent that randomly selects its action. Herein, we created our agents using the AI wolf server [4] released on the Artificial Intelligence Werewolf site<sup>2</sup>. We used data from 467 instances of the werewolf game with 223 villager wins and 224 werewolf wins as the action selection probability of the agent. We used K-fold cross-validation with  $K = 10$  to calculate the coincidence ratio; here, the coincidence ratio is the ratio that the agent’s highest-probability action in the situations defined in Sect. 5 above coincides with all human behaviors in the play logs in the same situation. Actions used to calculate the coincidence ratio include those of our proposed behavioral model, as shown in Table 4. For example, “I am a villager” is not used to calculate the coincidence ratio because it is not available in our proposed behavioral model. Furthermore, when the situation is as summarized in Table 5, the selectable actions regarding the attack target of a werewolf are shown in Table 6. The agent’s highest-probability action from among selectable actions is  $p_2$  of Table 6. We could thus investigate whether the agent’s action coincides with human behavior by comparing  $p_2$  with human behaviors obtained from play logs from the Werewolf BBS.

### 6.2 Results

Figure 1 shows the degree of coincidence between the actions determined by our behavioral model and human behavior. In this figure, green bars indicate the number of games. As the game can be finished in five days at the earliest, the number of games gradually decreases from the sixth day onward. The average degree of coincidence of our proposed model was 81.55%, whereas that of the random model was 33.73%. From Fig. 1, the agent based on our behavioral model of action selection probability behaved like a human substantially more so than the random agent. The number of executable actions of the agent, e.g., execute, attack, and guard, increased in the middle days of the game given the increase in the number of CO actions and players inspected by a seer. Therefore, as the game reaches its middle, the degree of coincidence of the random agent further decreases. Furthermore, the degree of coincidence of the agent based on our behavioral model of action selection probability decreased, too, but obtained higher values than that of the random agent.

In the final phase of the werewolf game, the coincidence ratio of the agent based on our behavioral model of action selection probability increased; however, the tendency that this degree of coincidence decreases was seen in days 8 and 9. This tendency is not seen in the random agent because the action selection

<sup>2</sup> <http://www.aiwolf.org/en/>.



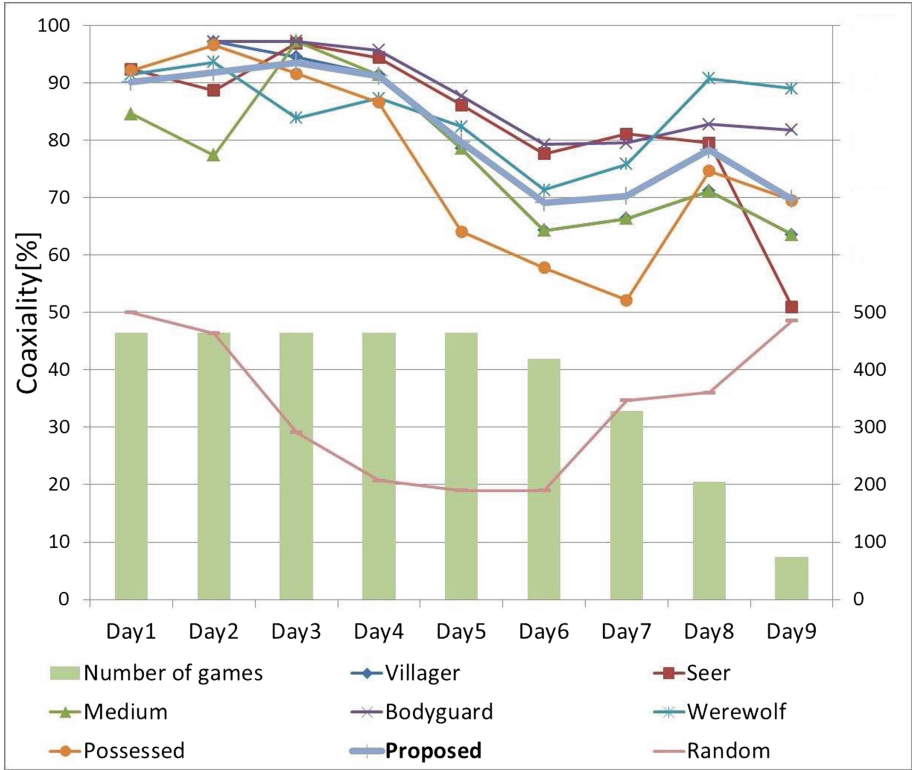


Fig. 1. Degree of coincidence (Color figure online)

probability cannot be obtained for day 9 given that the situation information is insufficient. There are only 74 games remaining on day 9.

### 6.3 Consideration

We conducted our simulation experiments 10,000 times on agents based on our behavioral model of action selection probability to investigate its influence on victory or defeat. In our simulation experiment, we prepared 15 agents based on our behavioral model of action selection probability; these agents played the werewolf game between the same agents. As noted above, we also conducted the same simulation experiments using the random agent to compare our results.

The winning percentage of our simulation experiments is shown in Table 7. We calculated the winning rate of actual games by using the data of the aforementioned 467 instances (with 223 villager wins and 224 werewolf wins). As summarized in the table, the obtained winning rate of agents based on our behavioral model of action selection probability was closer to the winning rate of the actual games when compared to that of the random agent; however, there is still a substantial difference between the winning percentage of actual games and that

**Table 7.** Winning percentage.

	Villager win	Werewolf win
Actual game	47.8%	52.2%
Proposed method	32.6%	67.4%
Random method	22.0%	78.0%

of the agents based on our behavioral model of action selection probability. This discrepancy may have occurred because our proposed behavioral model can distinguish players by their CO and decision information; however, if players express the same role and are given the same judgment, our proposed behavioral model cannot distinguish between those players. More specifically, given the situation summarized in Table 5, our behavioral model based on action selection probability identifies actions from Table 6 as well as each action’s selection probability; however, our behavioral model cannot distinguish a seer of  $p_1$  from another seer of  $p_1$ . Similarly, our behavioral model cannot distinguish player  $p_2$  from another player of  $p_2$ .

In actual gameplay, when a human player comparatively selects one player as a target of an action in the above situation, the human player considers discrepancies from past utterances, impressions of others, etc. Our future work aims to incorporate these aspects to construct a refined behavioral model that can distinguish between players in the same role.

## 7 Related Work

Related studies include Monopoly [6] and The Settlers of Catan [7, 8]. Both these games attempt to include communication in the gameplay via the computer; however, utterances of these games target negotiation, e.g., in the trading and exchange of properties, utilities, and the like. This only requires the ability to estimate the intentions of others. These studies are related to automated negotiation, which is widely studied in the field of multi-agent systems (e.g. [9, 10]). Conversely, in the werewolf game, the ability to persuade and earn credibility is as important as estimating the intentions of other parties through logical thinking.

Taylor investigated “The Resistance Game” wherein trust affects the game result in common with the werewolf game [11]. However, this study focuses on the game without communication among players.

There are a few existing studies on the werewolf game. Braveman [12] and Yao [13] both showed that the probability of a werewolf-side win,  $w(n, m)$ , is proportional to  $m/\sqrt{n}$ , where  $n$  is the number of players at the start of the game and  $m$  is the number of werewolves. Furthermore, Migdal showed the exact formula of probability  $w(n, m)$  [14].

In these studies, players with special abilities (e.g., seers) are not included, thus simplifying the mathematical modeling. The game is performed using only

villagers and werewolves; however, actual games include many more roles, as seen on the Werewolf BBS. We also note that there are substantial differences regarding the process and nature of the game when roles other than just villager and werewolf are included. For example, if the aforementioned Werewolf BBS data were applied to the expression provided by Migdal (assuming the roles with special abilities are lumped into the villager role, i.e.,  $n = 15$  and  $m = 3$ ), we obtain a werewolf winning rate of 97.1%; however, according to our research, the actual werewolf winning percentage on the Werewolf BBS was 52.2%.

There are studies that have focused on human behavior and the psychological aspects of playing a werewolf that used various features for determining whether a player is a werewolf. For example, there was a study that used each player's utterances, utterance lengths, and the number of interruptions [15]; a study that used hand and head movements [16]; and a study that used the number of words in each utterance [17] to determine whether a player was a werewolf. Furthermore, several audio–visual corpus containing dialogue data in the werewolf game were constructed to analyze group communication [18,19]. However, these studies do not focus on playing the werewolf game with a computer.

## 8 Conclusion

In this study, we constructed a behavioral model by obtaining behavioral information from play logs describing play between humans; our model identifies an action selection probability to realize an agent that can behave like humans. We first obtained two types of information, i.e., CO and decision information via regular expressions to obtain behavioral information, and then acquired information regarding the dead or alive state at each date, the role of each player, etc. Consequently, we obtained a precision of 98.4%, a recall of 98.4%, and an f-measure of 0.98 for CO information acquisition; for decision information, we obtained a precision of 83.2%, a recall of 73.1%, and an f-measure of 0.78. We constructed a behavioral model based on action selection probability using information acquired from play logs and conducted simulation experiments. Consequently, agents based on our behavioral model of action selection probability behaved like humans much more so than a random agent; however, action selection probabilities could not be obtained in some instances due to insufficient situation information. In future work, we aim to include more game data and work to distinguish between players that express themselves as the same role and are inspected by the same side.

**Acknowledgements.** We heartily thank Mr. Ninjin for allowing us to use the data of the Werewolf BBS. This study received a grant of JSPS Grants-in-aid for Scientific Research 15K12180.

## References

1. Ito, T., Kitasei, Y.: Proposal and implementation of “digital curling”. In: 2015 IEEE Conference on Computational Intelligence and Games (CIG), pp. 469–473 (2015)

2. Garcia-Sanchez, P., Tonda, A., Mora, A.M., Squillero, G., Merelo, J.J.: Towards automatic StarCraft strategy generation using genetic programming. In: 2015 IEEE Conference on Computational Intelligence and Games (CIG), pp. 284–291 (2015)
3. Osawa, H., Toriumi, F., Katagami, D., Shinoda, K., Inaba, M.: Designing protocol of werewolf game: protocol for inference and persuasion. In: The 24th Fuzzy, Artificial Intelligence, Neural Networks and Computational Intelligence (FAN 2014) (2014)
4. Toriumi, F., Kajiwara, K., Osawa, H., Inaba, M., Katagami, D., Shinoda, K.: Development of AI wolf platform: development of game ai by using collective intelligence. In: Digital Games Research Association JAPAN 2014 (DiGRA JAPAN 2014) (2015)
5. Katagami, D., Takaku, S., Inaba, M., Osawa, H., Shinoda, K., Nishino, J., Toriumi, F.: Investigation of the effects of nonverbal information on werewolf. In: 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 982–987 (2014)
6. Yasumura, Y., Oguchi, K., Nitta, K.: Negotiation strategy of agents in the Monopoly game. In: Proceedings 2001 IEEE International Symposium on Computational Intelligence in Robotics and Automation, pp. 277–281 (2001)
7. Guhe, M., Lascarides, A., O'Connor, K., Rieser, V.: Effects of belief and memory on strategic negotiation. In: Proceedings of the SEMDIAL, vol. 2013, pp. 82–91 (2013)
8. Guhe, M., Lascarides, A.: The effectiveness of persuasion in the settlers of catan. In: IEEE Conference on Computational Intelligence and Games (CIG), pp. 1–8. IEEE (2014)
9. Okumura, M., Fujita, K., Ito, T.: An implementation of collective collaboration support system based on automated multi-agent negotiation. In: Ito, T., Zhang, M., Robu, V., Matsuo, T. (eds.) Complex Automated Negotiations. SCI, vol. 435, pp. 125–141. Springer, Heidelberg (2013)
10. Lang, F., Fink, A., Brandt, T.: Design of automated negotiation mechanisms for decentralized heterogeneous machine scheduling. *Eur. J. Oper. Res.* **248**(1), 192–203 (2016)
11. Taylor, D.P.: Investigating approaches to ai for trust-based, multi-agent board games with imperfect information; with don eskridge’s “the resistance”, *Discovery, Invention & Application*, no. 1 (2014)
12. Braverman, M., Etesami, O., Mossel, E.: Mafia: a theoretical study of players and coalitions in a partial information environment. *Ann. Appl. Probab.* **18**(3), 825–846 (2008)
13. Yao, E.: A theoretical study of mafia games, arXiv preprint [arXiv:0804.0071](https://arxiv.org/abs/0804.0071) (2008)
14. Migdal, P.: A mathematical model of the Mafia game, arXiv preprint [arXiv:1009.1031](https://arxiv.org/abs/1009.1031) (2010)
15. Chittaranjan, G., Hung, H.: Are you awerewolf? detecting deceptive roles and outcomes in a conversational role-playing game. In: 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 5334–5337 (2010)
16. Xia, F., Wang, H., Huang, J.: Deception detection via blob motion pattern analysis. In: *Affective Computing and Intelligent Interaction*, pp. 727–728 (2007)
17. Zhou, L., Sung, Y.W.: Cues to deception in online Chinese groups. In: Hawaii international conference on system sciences. In: Proceedings of the 41st Annual, pp. 146–146 (2008)
18. Hung, H., Chittaranjan, G.: The idiap wolf corpus: exploring group behaviour in a competitive role-playing game. In: Proceedings of the International Conference on Multimedia, pp. 879–882. ACM (2010)
19. Prévot, L., Yao, Y., Gingold, A., Bel, B., Chan, K.Y.J.: Toward a scary comparative corpus: the werewolf spoken corpus. In: SEMDIAL 2015, pp. 204–205 (2015)