

# AI Wolf Contest — Development of Game AI Using Collective Intelligence —

Fujio Toriumi<sup>1(✉)</sup>, Hirotaka Osawa<sup>2</sup>, Michimasa Inaba<sup>3</sup>, Daisuke Katagami<sup>4</sup>,  
Kosuke Shinoda<sup>5</sup>, and Hitoshi Matsubara<sup>6</sup>

<sup>1</sup> The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan  
tori@sys.t.u-tokyo.ac.jp

<sup>2</sup> University of Tsukuba, Tsukuba, Japan

<sup>3</sup> Hiroshima City University, Hiroshima, Japan

<sup>4</sup> Tokyo Polytechnic University, Tokyo, Japan

<sup>5</sup> The University of Electro-Communications, Chofu, Japan

<sup>6</sup> Future University Hakodate, Hakodate, Japan

**Abstract.** In this study, we specify the design of an artificial intelligence (AI) player for a communication game called “Are You a Werewolf?” (AI Wolf). We present the Werewolf game as a standard game problem in the AI field. It is similar to game problems such as Chess, Shogi, Go, and Poker. The Werewolf game is a communication game that requires several AI technologies such as multi-agent coordination, intentional reading, and understanding of the theory of mind. Analyzing and solving the Werewolf game as a standard problem will provide useful results for our research field and its applications. Similar to the RoboCup project, the goal of this project is to determine new themes while creating a communicative AI player that can play the Werewolf game with humans. As an initial step, we designed a platform to develop a game-playing AI for a competition. First, we discuss the essential factors in Werewolf with reference to other studies. We then develop a platform for an AI game competition that uses simplified rules to support the development of AIs that can play Werewolf. The paper reports the process and analysis of the results of the competition.

## 1 Introduction

The development of an artificial intelligence (AI) player that can play a game with a human has been one of the main benchmarks in the AI field for researching intelligence and its requirements. In the field of complete-information games, such as Chess or Shogi, AIs have already defeated top human players. In 2016, an AI defeated a human in the last-remaining complete-information game, Go [1]. In the field of incomplete games, Texas Hold'em Poker is a conventional game that organizes competitions in the AI field [2]. In 2015, Bowling et al. solved the problem of two-player-limited Poker [3]. Further, action video games are starting to be used for evaluating AI in real-time situations [4].

Compared to previous game challenges, communication or communicative intelligence, which is commonly used in board and card games, has not been attempted. When users play board and card games, they also converse with other players. Furthermore,

some games are actually conducted through conversations, and these are referred to as communication games. Relatively few studies focus on the application of AI in such communication games.

“Are You a Werewolf?” is one of the most popular communication games. The cover story of the “Werewolf” game (also known as “Mafia”) is as follows. “It’s a story about a village. Werewolves have arrived who can change into and eat humans. The werewolves have the same form as humans during the day, and attack the villagers one-by-one every night. Fear, uncertainty, and doubt towards the werewolves begin to grow. The villagers decide that they must execute those who are suspected of being werewolves, one by one...”

The winner of the Werewolf game is decided solely through discussions. Consequently, game players must use their cognitive faculties to the full. In contrast to a perfect-information game, players hide considerable information. Every player attempts to determine the hidden information by using other players’ conversations and behaviors, while trying to hide his/her own information to accomplish the objective. The game highlights various problems that have not been addressed adequately in the area of AI, such as an asymmetric diversity of player information, persuasion as a method of earning confidence, and speculation as a method of detecting fabrication.

Therefore, we started a project to create an AI Werewolf (AI Wolf), which plays the Werewolf game in place of a human. In addition, there are several trials for improving AI with game competitions such as Lemonade Stand Game competition [5] and Annual Computer Poker competition [6].

This is a comprehensive project, which aims at the development of not only a game-playing algorithm but also virtual agents and real robots. Many tasks must be solved to achieve the stated objective. Similar to the RoboCup [7] approach in robotics, to solve these tasks, we employ a collective-intelligence approach, which uses competition to improve each player’s algorithm. A common platform is indispensable when implementing a collective-intelligence approach. In this paper, we describe the outline of the Werewolf platform for AI (AI Wolf Platform) that we developed as an open-source project. We plan to organize a tournament of AI Wolf in which researchers from various backgrounds can participate freely, with the aim of realizing collective intelligence with the participating researchers.

Section 2 consists of the related studies about incomplete information games and Sect. 3 provides the overview of the werewolf games. Next, the AI Wolf project is defined in Sect. 4, and we describe our analysis for the first competition in Sect. 5. Finally, Sect. 6 provides the conclusion and future proposals.

## 2 Challenges for Incomplete-Information Games

The development of a game-playing agent has been a challenge from the beginning of AI research [8]. Several two-player board games with perfect information, such as Checkers, Othello, Chess, and Go, have been used for trials by applying several new algorithms [9, 10]. In these games, all information is observable by both players. An AI system must only handle the condition of the board and does not need to determine a competitor’s thought processes.

Further, there are several unsolved games in incomplete information game fields. Card games have information that cannot be observed by other players [11]. This is also an important field in AI research. Poker is one of the best-known examples, on which several theoretical analyses have been conducted [12]. Other games, including Bridge and the two-player version of Dou Zi Zhu (a popular game in China), have also been studied [13, 14]. Compared to these games, the Werewolf game requires intelligence to estimate the roles and internal states of other players. Although their information cannot be observed by other players, each player's role in the aforementioned games is determined before the game starts and is known to all players. In contrast, a player's role in the Werewolf game is hidden from the other players and is only revealed at the end of the game. This type of situation requires more intelligence because each player (especially a villager) needs to hold multiple world models for the other players' actions. It also suggests that a stable strategy does not exist because if some action suggests that a player supports the villagers, a werewolf will mimic this action. Inaba analyzed the change in the theory in the online werewolf game called "werewolf bulletin board system (BBS)" [15] for 10 years. In addition, this game requires persuasion of other players. This type of intelligence requires two levels of the Theory of Mind: the expectation of other players' expectations [16]. All these considerations suggest that research on the Werewolf game will lead to several new findings in the field of AI.

## 2.1 Studies on the Werewolf Game

There have been some studies on the Werewolf (or Mafia) game, including a mathematical analysis [17, 18]. In addition, some researchers have attempted to detect a player's role by using the length of utterances and the number of interruptions of a speaker [19], nonverbal information [20], hand and head motions [21], and the words used in discussions [22].

Some researchers have used the Werewolf game as a study for human-agent interactions. Aylett et al. [23] applied the Werewolf game for educating children on cultural sensitivity. Katagami et al. [20] investigated the effect of nonverbal information in the Werewolf game. However, studies have not attempted to develop playing agents.

To realize Werewolf playing agents, many tasks must be solved. These include the asymmetric diversity of player information, persuasion as a means of earning confidence, and speculation to detect fabrication. These tasks are not generally considered in the field of AI agents.

## 3 AI Wolf Project

### 3.1 What Is "Are You a Werewolf?"

#### Overview

Werewolf is a popular party game played worldwide. Werewolf card sets include "Are You a Werewolf?" and "Lupus in Tabula." The game is still played around the world. Additionally, "Mafia" has an identical game structure but with a much less magic-based theme. "Are you a Werewolf?" is a party game that models a conflict between an

informed minority and an uninformed majority. Initially, each player is secretly assigned a role affiliated with one of these teams. There are two phases: night and day. At night, the werewolves “attack” the townsfolk. During the day, surviving players discuss the elimination of a werewolf by voting. The objective of the werewolves is to kill off all the villagers without being killed themselves. The objective of the townsfolk is to ascertain who the werewolves are and to kill them.

There are two techniques for playing Werewolf. The first includes face-to-face play by using the game cards described earlier. The other is to play online using web applications, or a BBS-type platform. For example, large BBS services exist in Japan for playing Werewolf. In fact, there are more than a thousand logs of Werewolf games. Many players still play Werewolf on a BBS. Moreover, some academic studies make use of the BBS game logs. Developing physical robots that can play Werewolf face-to-face is one objective of our project; however, many problems must be solved. Consequently, in this paper, we use a simplified representation of the essence of the game based on a BBS-type Werewolf.

### **Game Procedures**

The roles of all players are allocated randomly. Players are divided into two teams, townsfolk and werewolf teams, according to their roles and the method of winning of their teams. The victory condition for the townsfolk is to kill all the werewolves. For the werewolves, the victory condition is to kill humans such that they become equal or fewer in number to the werewolves. A player fundamentally cannot know the other players’ roles because the allocated roles are unpublicized. A basic course of action for the townsfolk players is to discover werewolves through conversation because they do not know who the werewolves are. In contrast, the werewolf players know who the werewolves are. Therefore, a basic course of action for the werewolf players is to engage in various cooperative maneuvering, without the townsfolk knowing about their roles.

The game proceeds in alternating phases of day and night. During the day, all players discuss who the werewolves are. Simultaneously, players who have special abilities (which we discuss later) lead discussions that produce advantages for their respective teams by using the information derived from their abilities. After a certain period, players execute one player who is suspected of being a werewolf, as chosen by majority voting. The executed player then leaves the game and cannot play. During the night, werewolf players can attack a townsfolk team player. The attacked player is killed and is eliminated from the game. In addition, players who have special abilities can use those abilities during the night phase. The day and night phases alternate until the winning conditions are met.

Townsfolk players must be able to detect a werewolf player’s lie. In addition, persuading other players by using the information obtained through their special abilities is important. Furthermore, a crucially important point for the werewolf team is to manipulate the discussion to the team’s advantage. Occasionally, they must impersonate a role and obfuscate the conditions and evidence.

### **Roles of Players**

There are many variations of the rules and roles of the Werewolf game. Therefore, we use the following basic set of roles for simplification.

- Villager: Townsfolk team. A character in this role has no special ability.
- Werewolf: Werewolf team. Werewolves can attack one townsfolk player during each night phase. They all decide on a single player to attack together with vote, and zero or one villager dies each night. BBS-type game also allows werewolves to talk with each other simultaneously during the day, and we used the same rules in this AI game.
- Seer: Townsfolk team. A seer can inspect a player in every night phase to ascertain whether or not a player is a werewolf.
- Bodyguard: Townsfolk team. A bodyguard can choose a player in every night phase and protect the player against an attack by a werewolf.
- Medium: Townsfolk team. A medium can ascertain whether a player who was executed during the previous day phase was a werewolf.
- Possessed: Werewolf team. Werewolves do not know who is a possessed player. The possessed have no special ability. This role secretly cooperates with werewolves because a werewolf-team victory is also regarded as a victory for possessed players.

### 3.2 Roadmap of AI Wolf Project

We plan to create AI agents that can play the Werewolf game [9]. It is an incomplete-information game. In addition, the Werewolf game is conducted solely through discussion, and players must use their cognitive faculties completely to win. The symbolization of the Werewolf game is difficult compared to other incomplete-information games such as Poker. This feature requires a different approach than other incomplete-game challenges.

An AI agent requires multiple research areas, such as analyzing the human playing Werewolf, natural language processing, agent technology, and human-agent interaction. Our project consists of not only a sole project team but also of multiple research teams. Figure 1 explains the milestones of the project, the keystone of which is a Werewolf intelligence competition (WIC) that gathers the collective intelligence of people in a program.

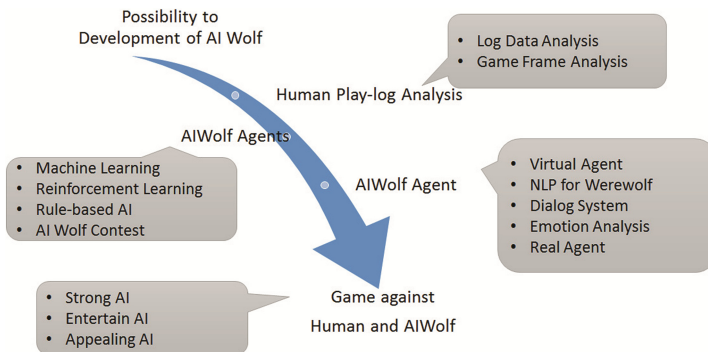


Fig. 1. Project plan of WIC

## 4 AI Wolf Platform

### 4.1 Architecture of the AI Wolf Platform

We have been developing the AI Wolf Platform, which is intended to function as an apparatus for evolving AI Wolf agent Game Player algorithms through collective intelligence. The platform consists of the game server and game-player agents (as shown schematically in Fig. 2). These agents connect to the server and play the Werewolf game. Therefore, this platform is built on the client-server architecture. The game server performs the role of game moderator. Moreover, the server controls the network communication between the agents and itself and maintains a log of the games. Game-player agents communicate with the game server via TCP/IP or an internal function-call API. By using the TCP/IP connection, developers can play against other wired player agents. In addition, by using the internal function call, developers can conduct high-speed simulations. The AI Wolf Platform has a communication protocol API between the server and clients. This API is an abstraction layer for the game server and player agents. It facilitates parsing by restricting communication to a specific content format.

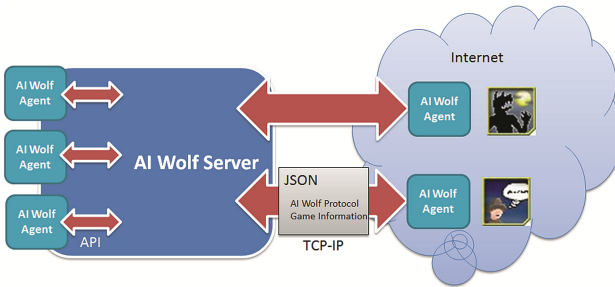


Fig. 2. AI Wolf platform architecture

The game server library is offered by Java. Agent-building libraries are offered by Java, .NET Framework, and Python. Agent programmers need not be concerned about communication between server and client because the communication protocol is wrapped by a library. Agent programmers simply implement the Player interface. All agent classes work by an event-driven method. The server asks clients for their behavior, and clients reply accordingly.

### 4.2 Development of an AI Wolf Agent

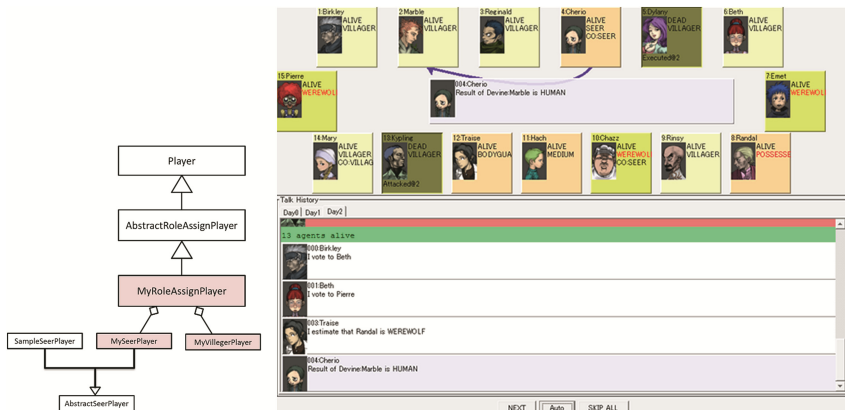
Each agent acts through event-driven systems. The game server sends a request and agents return a response as an action in the Werewolf game. Table 1 shows the requests that are sent from the AI Wolf Server to the agents. Therefore, a game-agent developer must consider only how agents should act when each request arrives. In summary, the developer must implement some method that corresponds to each request.

**Table 1.** Requests from the AI Wolf Server

Request	Agent action	Reply
Initialize	Initialize for game start	–
DailyInitialize	Initialize for day Start	–
Finish	Finish the game	–
Name	Return name of the agent	Name
Role	Return the role of an agent	Role
Talk	Talk to other agents	Talk
Whisper	Talk to other werewolves	Whisper
Vote	Choose an agent to be voted	Agent
Divine	Choose an agent to be divined	Agent
Guard	Choose an agent to be guarded	Agent </td
Attack	Choose an agent to be attacked	Agent

In the Werewolf game, an agent should change its behavior pattern depending on its role. The possible requests differ for each role.

To simplify the accommodation of different roles, the AI Wolf agent library contains the class of AbstractRolePlayer (as shown at the left in Fig. 3). When the developers implement a role, they program the AbstractPlayer class (e.g., in the case of Seer, it would be AbstractSeerPlayer) and assign its class as the function MyRoleAssignPlayer. For example, if we wish to implement a Seer player, we program AbstractSeerPlayer and assign it to MyRoleAssignPlayer, whereupon the agent can act as a seer player. This AbstractRoleAssignPlayer has default behaviors for all roles. The developer should not have to create all the role-behavior algorithms but can instead use default algorithms (or those of other developers) from an early stage.



**Fig. 3.** Class diagram of AbstractRolePlayer (left) and GUI log viewer (right)

Furthermore, we implemented a GUI log viewer (shown at the right in Fig. 3) to help with program debugging. It can be used not only for showing the behavior of agents but also for interactive debugging when programming an agent's behavior.

### 4.3 AI Wolf Protocol

During the game, agents communicate with other agents using the AI Wolf Protocol, which is a shortened communication protocol designed for AI Wolf. This communication protocol is determined by referring to frequent utterances used in Werewolf BBS. The Werewolf BBS allows a limited number of communications (20 per day for the villagers, and 30 per day for werewolves); this limitation causes shortened symbols. For example, the expressing of a role is called "coming out" (divulgence), which is shortened to "CO." As such, CO and other designators are used distinctively; we applied this difference to our protocol.

The current version of AI Wolf Platform employs a simple protocol as the first step of the project. This simple protocol permits only limited utterances, such as "I declare as seer" and "I suspect that he is a werewolf." We evaluated the Werewolf BBS logs, in which 50% of the utterances are represented through 10 protocols. Hence, each agent can use the following 10 communication protocols as explained:

- estimate(Agent, Role)
  - An agent expresses its suspicion that [Agent] is [Role].
- comingout(Agent, Role)
  - The agent asserts that [Agent] is [Role].
- divined(Agent, Species)
  - The agent (implicated as a seer) gives the divined result that [Agent] is [Species (human or werewolf)]
- inquested(Agent, Species)
  - The agent (implicated as a medium) gives the inquested (investigated) result that the executed [Agent] is [Species (human or werewolf)]
- guarded(Agent)
  - The agent (implicated as a bodyguard) gives the result that [Agent] is protected.
- vote(Agent)
  - The agent claims that a player will select [Agent] for the execution vote
- agree(day, id)
  - The agent agrees with someone's statement at statement number [id] on [day].
- disagree(day, id)
  - The agent disagrees with someone's statement at statement number [id] on [day].
- skip()
  - The agent skips its turn to talk, and waits for the next turn. That is, the agent waits to listen to an opponent's talk and wishes to continue the discussion.
- over()
  - The agent skips its turn to talk, waits for the next turn, and agrees to finish its discussion the same day.



To ease the development of AI Wolf agents, the platform provides an utterance factory and parser for the protocol.

## 5 The First WIC

We organized the first WIC at the Computer Entertainment Developers Conference (CEDEC) on August 27, 2015. CEDEC is one of the biggest domestic conferences in Japan for video-game competitions, and is being organized since the last 17 years. More than 30,000 people participate in the conference. Representatives from academic institutions and video-game companies attend and exchange their findings. Moreover, several international research sessions are organized. Thus, we considered that this conference would be a good forum to evaluate our approach.

### 5.1 Rules of the Competition

We organized preliminary and final competitions. Both competitions were staged according to a BBS-type Werewolf game. Fifteen agents joined one game set, and roles listed in Table 2 were assigned to each agent. One set comprised 100 games, and agents was the same in each set but with different assigned roles. Each agent in the winning team received one point in each game.

**Table 2.** Roles and agents

Role	Count	Side
Villager	8	Villager
Seer	1	Villager
Medium	1	Villager
Bodyguard	1	Villager
Werewolf	3	Werewolf
Possessed	1	Werewolf

In the final competition, 1,124,890 games were played and 15 agents were assigned their roles randomly in the games. All the games were ranked.

### 5.2 Participants

Each participating team could submit one agent program. 78 teams joined the competition, and 45 teams submitted programs. Seven agents were rejected because of errors; hence, 38 agents joined the preliminary competition.

Table 3 shows the fraction of students in the competition. More than half of the participants comprised students from universities and other educational organizations. This result suggests that these participants not only joined in with the programming competition, but also focused on the research associated with the Werewolf game (several students also published research on the Werewolf game after the competition).

However, the lower rate of students in the final competition suggests that the professionals, including video programmers, were more proficient than the students.

**Table 3.** Participating teams

	Total	Students	Fraction of Students
Registered	78	42	0.53
Preliminary	38	24	0.63
Final	15	7	0.47

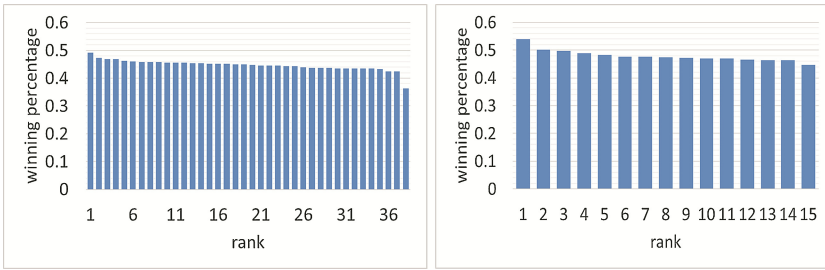
Figure 4 shows some of the participants during the final competition at CEDEC 2015. More than 200 participants took part in the final competition, and our results were reported through at least five media outlets. In the CEDEC 2015 session, we selected one example from the final competition, and participants explained their algorithms to each other.



**Fig. 4.** Final competition in CEDEC 2015

### 5.3 Results of the Competition

The left of Fig. 5 shows the success rates of all 38 teams who took part in the preliminary competition. According to this result, although high-ranking agents generally have higher success rates, most agents have rates that are approximately the same. This may be because Werewolf is a multiple-player game. As such, each agent's contribution toward a win is lower than it would be in a single-player game. The rate of the top-ranked agent is 0.4915, whereas the rate of the bottom-ranked agent is 0.3629. This represents a 13% difference between the strengths of agents. There is no significant difference between the agents ranked 15th and 16th, and participants in this rank border are assigned by luck. We need to improve our rules for the next competition to reflect significant differences at the borders.



**Fig. 5.** Success rate of agents in preliminary competition (left) and in final competition (right)

The right of Fig. 5 shows the success rates of all 15 agents who took part in the final competition. We statistically analyzed the difference between two odd ratios. The result suggests that the five top-ranked agents are significantly stronger than the other 10 agents.

#### 5.4 Analysis of the Final Competition

Table 4 shows the types of roles in which an agent was strongest. The results suggest that the top-ranked agents are strong in nearly every role.

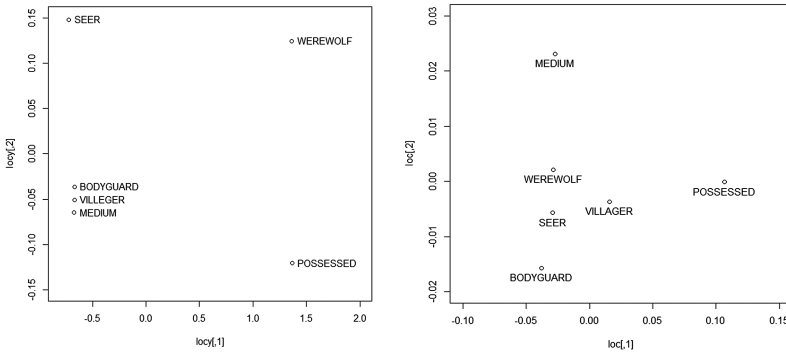
**Table 4.** Ranking for each role

Final rank	Villager rank	Seer rank	Medium rank	Bodyguard rank	Werewolf rank	Possessed rank
1	1	1	7	1	1	1
2	4	2	8	2	3	4
3	3	3	3	8	2	14
4	2	11	12	4	4	5
5	10	7	1	5	5	3
6	9	4	9	10	6	8
7	11	9	4	6	12	2
8	6	10	15	3	10	6
9	14	5	10	7	7	12
10	5	14	2	15	8	11
11	13	6	11	11	13	7
12	8	13	5	12	11	10
13	12	8	13	13	9	15
14	7	15	6	9	14	9
15	15	12	14	14	15	13

First, we calculated the total success rate for each role and the difference between square errors of these rates. Next, we plotted the results using multidimensional scaling. We calculated the distance  $D_{kl}$  between points  $k$  and  $l$  according to the following equation:

$$D_{kl} = \sum_i (x_{ik} - x_{il})^2$$

The value of  $x_{ik}$  represents the success rate of role  $k$  for agent  $i$ . Figure 6 (left/right) shows the data from the preliminary/final competition. The success rates for Villager, Medium, and Bodyguard are relatively close, whereas those for Seer, Werewolf, and Possessed are distinctly different. The Seer, Werewolf, and Possessed roles require more specific skills, thus explaining the large distances in the plot in Fig. 6. In the final competition, the Medium and Possessed roles showed different behaviors than in the previous competition. We speculate that players in the final competition wrote more intelligent code than those in the preliminary competition. For a Medium, the highest and lowest scoring agents showed very slight difference, indicating that a Medium does not contribute much toward winning: a result similar to those obtained through statistical analyses of online Werewolf games in Japan. Moreover, some high-ranking players score lower than the low-ranking players in the Possessed role. This suggests that the Possessed role requires certain unique features than the other roles. We speculate that this apparent fact helped players with their programming-resource management from one competition to the next.



**Fig. 6.** Plots for success-rate distance of each role (left: preliminary competition; right: final competition).

Lastly, we evaluated the difference between the success rates of agents who did and did not reach the final according to each role (Fig. 7). In all roles, finalists are stronger by a significant difference. In particular, the differences range between 4% and 5% for both Seer and Werewolf roles. This fact may be helpful in suggesting where programmers could best focus AI in the Werewolf game.

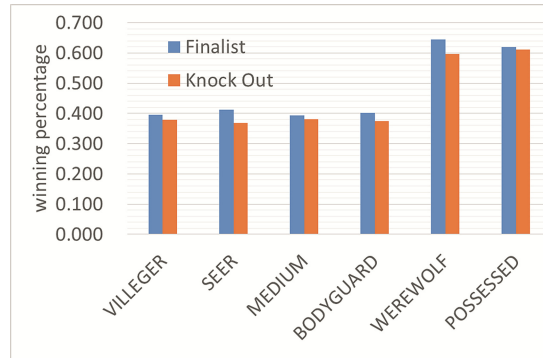


Fig. 7. Success rates of preliminary and finalist agents in each role

## 5.5 Discussion

There are several trends that are observed in the participating agents.

1. Strong agents tend to be strong in any role.
2. Seer, Werewolf, or Possessed scores might differ from their real ability in the preliminary competition.
3. Although higher-ranked agents tend to score higher, the success rates of Medium and Possessed roles differ from the other roles.
4. All agents in the final competition are significantly stronger than unsuccessful agents in any role. The difference is especially clear for both Seer and Werewolf.

Our findings suggest that these competition trials facilitate a collective-intelligence approach, the findings of which contribute to the analysis of Werewolf games. However, there are still some agents with anomalous behavior, even in the final competition. These difficulties may improve in future challenges. For further evaluation, we want to reflect evaluation methods for other competitions [24, 25].

The first competition was reported by several media outlets (NIKKEI, Game Watch, GPara.com, ASCII/Digital). We believe that our challenge made a good impact on society, and the notion of “lying AIs” stimulated discussions on the role of AI in modern society.

## 6 Conclusion and Future Work

This paper summarized the AI Wolf project, its competition server, and results from the first WIC. Thirty-eight agents participated in the competition, with 15 agents participating in the finals. The top agent UDON was significantly stronger than all the other agents. The results of the study support the assertion that a competition facilitates in achieving collective intelligence.

The analysis of the relationship between roles and success rates reveals several role-dependent features. These agent sources are available on an open-source basis on our

site, and agents who outsmart these programs are likely to win in the next competition. This step “evolves” agent strategy, and we expect that this ecosystem will produce new findings for the AI field about the meaning of communication. Our source codes are completely open. The server codes are available on GitHub (<https://github.com/aiwolf/>), and the final agent source codes are available on the AI Wolf project site (<http://aiwolf.org/>). We are now planning a second competition for CEDEC 2016.

In the future, we want to evaluate the value of entertainment of the Werewolf game. Werewolf is a type of party game. People play party games for not only winning and losing but also enjoying the game and its communication element. Therefore, an AI for Werewolf that can “entertain players” must be established. We assume that if the agents are sufficiently strong to defeat all the humans without any entertainment, most people will avoid playing with them. In addition, we want to attempt to determine the source of pleasure in competitive games. Sometimes, the value of entertainment can be understood from a game’s rules. However, in many cases the action of opponents becomes a factor in determining the pleasure derived from a game. Therefore, to understand “why games entertain players,” it is necessary to consider the interaction between the opponent’s behavior and the game’s rules. In certain situations, players may enjoy a game. Such situations include competitive games between human players. In such a case, the purpose of AI can be assumed to be the generation of those pleasant situations. Therefore, the AI must act to change the game environment.

**Acknowledgements.** This work was partially supported by Hayao Nakayama Foundation for Science & Technology and Culture, Foundation for Fusion of Science and Technology, and JSPS KAKENHI Grant Number 26118006. We also want to say thanks for Computer Entertainment Developers Conference and Japan Society for Artificial Intelligence.

## References

1. Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D.: Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016)
2. Genesereth, M., Love, N., Pell, B.: General game playing: overview of the AAAI competition. *AI Mag.* **26**, 1–16 (2005)
3. Bowling, M., Burch, N., Johanson, M., Tammelin, O.: Heads-up limit hold’em poker is solved. *Science* **347**, 145–149 (2015)
4. Karakovskiy, S., Togelius, J.: The Mario AI benchmark and competitions. *IEEE Trans. Comput. Intell. AI Games* **4**, 55–67 (2012)
5. Zinkevich, M.A., Bowling, M., Wunder, M.: The lemonade stand game competition. *ACM SIGecom Exchanges* **10**, 35–38 (2011)
6. ACPC: Computer Poker Competition. [www.computerpokercompetition.org/](http://www.computerpokercompetition.org/)
7. Kitano, H., Asada, M., Kuniyoshi, Y., Noda, I., Osawa, E.: RoboCup: the robot world cup initiative. In: *Proceedings of the First International Conference on Autonomous Agents - AGENTS 1997*, pp. 340–347. ACM Press, New York (1997)

8. Abramson, B.: Control strategies for two-player games. *ACM Comput. Surv.* **21**, 137–161 (1989)
9. Krawiec, K., Szubert, M.G.: Learning N-tuple networks for othello by coevolutionary gradient search. In: *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation - GECCO 2011*, pp. 355–362. ACM Press, New York (2011)
10. Gelly, S., Kocsis, L., Schoenauer, M., Sebag, M., Silver, D., Szepesvári, C., Teytaud, O.: The grand challenge of computer Go. *Commun. ACM* **55**, 106–113 (2012)
11. Ganzfried, S., Sandholm, T.: Game theory-based opponent modeling in large imperfect-information games. In: *International Conference on Autonomous Agents and Multiagent Systems*, pp. 533–540. International Foundation for Autonomous Agents and Multiagent Systems (2011)
12. Billings, D., Burch, N., Davidson, A., Holte, R., Schaeffer, J., Schauenberg, T., Szafron, D.: Approximating game-theoretic optimal strategies for full-scale poker. In: *International Joint Conference on Artificial Intelligence*, pp. 661–668 (2003)
13. Ginsberg, M.L.: GIB: imperfect information in a computationally challenging game. *J. Artif. Intell. Res.* **14**, 303–358 (2001)
14. Whitehouse, D., Powley, E.J., Cowling, P.I.: Determinization and information set Monte Carlo tree search for the card game Dou Di Zhu. In: *2011 IEEE Conference on Computational Intelligence and Games (CIG 2011)*, pp. 87–94. IEEE (2011)
15. Inaba, M.I., Toriumi, F.U., Takahashi, K.E.: The statistical analysis of werewolf game data. In: *Proceedings of Game Programming Workshop 2012*, pp. 144–147 (2012) (in Japanese)
16. Dias, J., Reis, H., Paiva, A.: Lie to me: virtual agents that lie. In: *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 1211–1212 (2013)
17. Braverman, M., Etesami, O., Mossel, E.: Mafia: a theoretical study of players and coalitions in a partial information environment. *Ann. Appl. Probab.* **18**, 825–846 (2008)
18. Yao, E.: A theoretical study of Mafia games (2008). Arxiv Prepr. [arXiv:0804.0071](https://arxiv.org/abs/0804.0071)
19. Chittaranjan, G., Hung, H.: Are you a werewolf? Detecting deceptive roles and outcomes in a conversational role-playing game. In: *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 5334–5337. IEEE (2010)
20. Katagami, D., Kanazawa, M., Toriumi, F., Osawa, H., Inaba, M., Shinoda, K.: Movement design of a life-like agent for the werewolf game. In: *IEEE International Conference on Fuzzy Systems*, pp. 982–987 (2015)
21. Xia, F., Wang, H., Huang, J.: Deception detection via blob motion pattern analysis. In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) *ACII 2007 Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction*, pp. 727–728. Springer, Heidelberg (2007)
22. Zhou, L., Sung, Y.: Cues to deception in online chinese groups. In: *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, p. 146. IEEE (2008)
23. Aylett, R., Hall, L., Tazzyman, S., Endrass, B., André, E., Ritter, C., Nazir, A., Paiva, A., Höfstedt, G., Kappas, A.: Werewolves, cheats, and cultural sensitivity. In: *Autonomous Agents and Multi-Agent Systems*, pp. 1085–1092. International Foundation for Autonomous Agents and Multiagent Systems (2014)
24. White, M., Bowling, M.: Learning a value analysis tool for agent evaluation. In: *IJCAI 2009 Proceedings of the 21st International Joint Conference on Artificial Intelligence*, pp. 1976–1981 (2009)
25. Davidson, J., Archibald, C., Bowling, M.: Baseline: practical control variates for agent evaluation in zero-sum domains. In: *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems*, pp. 1005–1012 (2013)