

Generating a Werewolf Game Log Digest of Inferring Each Player's Role

Youchao Lin, Mizuho Baba

Graduate School of Systems and Information Engineering,
University of Tsukuba,
1-1-1 Tennodai, Tsukuba, 305-8573, Japan

Takehito Utsuro

Faculty of Engineering, Information and Systems,
University of Tsukuba,
1-1-1 Tennodai, Tsukuba, 305-8573, Japan

Abstract—While playing the communication game “Are You a Werewolf”, a player always guesses other players’ roles through discussions, based on one’s own role and other players’ crucial utterances. The underlying goal of this paper is to construct an agent that can analyze the participating players’ utterances and play the werewolf game as if it is a human. For the first step of this underlying goal, given a specific player participating in the wolf game, this paper studies how to generate a digest of inferring other players’ roles from the viewpoint of the given specific player. In this inference process, we regard the werewolf game rules as well as certain common sense as inference rules. Then, we develop a set of inference rules and apply them to infer the participating players’ roles from a real werewolf game log.

I. INTRODUCTION

Werewolf is a party game created in the USSR in 1986. It models a conflict between an informed minority, the werewolf, and an uninformed majority, the villager. The werewolf game has been popular in many countries including Japan. In Japan, particularly, not only the game itself, but several other activities such as “Werewolf TLPT” (Werewolf: the live playing theater)¹, a improvisation where the actors and actresses play the werewolf game, and a TV variety show² where comedians, actors, and actresses play the werewolf game.

In the research community of artificial intelligence, it has been well known that the werewolf game is one of games with imperfect information where certain information are hidden from some players. This situation is quite contrary to games with perfect information such as chess, shogi, and go, where it is known that computer programs won a human champion^{3,4,5}. In the Japanese research community of artificial intelligence, the werewolf game has been considered to be used as one of standard problems to evaluate the performance of general artificial intelligence since 2014 [1]. Also, research activities aiming at developing a computer agent program which participates in the werewolf game has started and the first competition of the AIWolf (artificial intelligence based werewolf)⁶ was held in August 2015 [2].

¹<http://7th-castle.com/jinrou/index.php> (in Japanese)

²<http://www.fujitv.co.jp/jinroh/index.html> (in Japanese)

³<http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/deepblue/>

⁴<http://www.shogi.or.jp/kisen/denou/> (in Japanese)

⁵<https://www.deepmind.com/alpha-go.html>

⁶<http://cedec.cesa.or.jp/2015/session/AC/7649.html> (in Japanese)

However, in those previous studies aiming at developing a computer agent program which participates in the werewolf game, research issues that are closely related to natural language processing and knowledge processing research have not been studied extensively. Those higher level research issues should include, e.g., understanding natural language conversations among the participating players, inferring each player’s roles considering the contents of their conversations, and deciding the player to be attacked or executed based on high level inference.

Considering the underlying goal of constructing an agent that can analyze the participating players’ utterances and play the werewolf game as if it is a human, as the first step, given a specific player participating in the wolf game, this paper studies how to generate a digest of inferring other players’ roles from the viewpoint of the given specific player. In this inference process, we regard the werewolf game rules as well as certain common sense as inference rules. Then, we develop a set of inference rules and apply them to infer the participating players’ roles from a real werewolf game log.

II. WEREWOLF GAME

In the werewolf game, each player is given a role and all the players are divided into one of the werewolf side and the villager side. Then, players of the both sides aim at winning the game. The werewolf side attacks one player of the villager side per day, while the villager side tries to execute one werewolf per day through arguments and votes. The players on the villager side do not know each player’s identity of being a werewolf or a human, while those on the werewolf side know those identifies. The werewolf side tries to make the players on the villager side vote themselves to be executed through misleading arguments by providing false information. Table I shows a typical case the list of roles of the both sides of the werewolf game with 16 players. Among those roles, the role of the possessed is on the werewolf side and the possessed wins when the werewolf side wins, although the seer divines the possessed to be a human, the medium declares the possessed to be a human as the result of the act of the medium, and the possessed is counted as a human when one survives.

Table II and Table III also list the rules and common sense of the werewolf game. The players are usually requested to follow those rules listed in Table II, while they are just assumed

TABLE I
ROLES IN THE WEREWOLF GAME (FOR 16 PLAYERS)

side	player type when counting the survivors	role	description	# of players
villager	human	villager	A human who does not have any special skill.	7
		seer	A human who belongs to the villager side. Every night, the seer can choose one player and learn whether the player is "werewolf" or "human". Learning the result, the seer can tell it to other players.	1
		medium	A human who belongs to the villager side. The medium can learn whether the player who was voted to be executed on the previous day is "werewolf" or "human". Learning the result, the medium can tell it to other players.	1
		bodyguard	A human who belongs to the villager side. Every night the bodyguard can choose one player except the bodyguard oneself to defend so that the chosen player can avoid being attacked by the werewolves. However, the bodyguard can not learn whether the player chosen to be defended was actually attacked or not.	1
		freemasons	A human who belongs to the villager side. The freemasons know each other's role.	2
werewolf	werewolf	possessed	A human who belongs to the werewolf side. The possessed wins when the the werewolf side wins. However, the possessed and werewolves do not know each others' roles.	1
		werewolf	Every night the werewolves choose one player on the villager side to be killed. The werewolves know each others' role and can communicate through a channel that are available only to the werewolves.	3
total			—	16

TABLE II
RULES OF THE WEREWOLF GAME

1	The number of the players for each of the roles of the seer, the medium, the bodyguard, and the possessed is one.
2	The number of the players for the role of the freemasons is two and the two freemasons know each others' role.
3	The werewolves know each others role.
4	The werewolves can not attack themselves.
5	When the number of the werewolves is larger than that of humans, the werewolf side wins.
6	When all the werewolves are executed, the villager side wins.

TABLE III
COMMON SENSE OF THE WEREWOLF GAME

1	The content of the utterances by the villagers, the seer, the medium, the bodyguard, and the freemasons do not conflict with the truth.
2	The seer / the medium reveal themselves as a seer / a medium.
3	The content of the utterances by the werewolves and the possessed may conflict with the truth.

to be follow those common sense listed in Table III. Those common sense are considered to be a kind of conventional strategies that are recommended to adopt so as to raise the winning rates of both the villager and the werewolf sides.

III. WEREWOLF GAME LOG

In this paper, as the werewolf game log data, we use that of WolfBBS⁷, which is a werewolf game site on the Internet, where the participating players communicate with each other with a character-based text input communication channel. This werewolf game site keeps the record of the text data of the previous werewolf game log and makes them publicly

⁷<http://ninjin002.x0.com/wolff/> (in Japanese)

available. As the result of our preliminary analysis on the text game log data, we confirm that the utterances and the actions that have strong influences on inferring each player's role can be extracted with sufficiently high precision⁸.

IV. WEREWOLF GAME LOG DIGEST OF INFERRING EACH PLAYER'S ROLE

This section introduces the fundamental idea of the werewolf game log digest we propose in this paper, where the most important function of the werewolf game log digest is to infer each player's role based on the utterances as well as the actions of all the players. The utterances and the actions which have strong influences on inferring each player's role in this paper are those eight types (more specifically, eight types of utterances and three types of actions) listed below:

- revealing oneself as the role of the seer, the medium, the bodyguard, and the freemasons (four types of utterances),
- the acts of the divination and the medium and declaring the results of the divination and the medium (two types of actions and two types of utterances),
- the act of bodyguard and then declaring whom the bodyguard defended (one type of actions and one type of utterances),
- one freemason's revealing the identify of the other freemason (one type of utterances).

Then, along with those eight types of utterances and actions as well as the rules and common sense of the werewolf game

⁸In the evaluation we present in section VI-B, we manually extract the utterances and the actions that have strong influences on inferring each player's role.

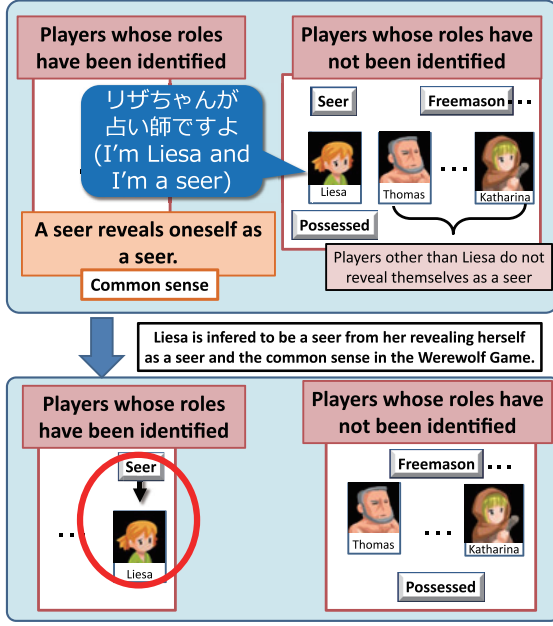


Fig. 1. An Example of Inferring Players' Roles

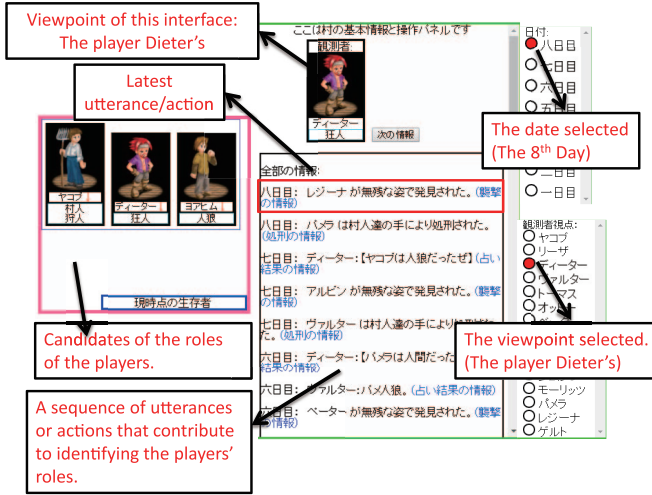


Fig. 2. A Snapshot of the Interface of Werewolf Game Log

shown in Table II and Table III, each player's role is inferred⁹

Fig. 1 shows a rough example of inferring a player's role based on the utterances, the actions, the rules, and the common sense of the werewolf game. In Fig. 1, we suppose that no player other than Liesa reveals oneself as the role of the seer on or before the third day. Then, based on this information as well as the common sense with ID=2 in Table III (the seer

⁹In addition to those rules and common sense, in this paper, we assume the following conventions: (i) The real seer, the possessed, and the werewolves reveal oneself as the role of the seer on or before the third day. The real medium, the possessed, and the werewolves reveal oneself as the role of the medium on or before the third day. (ii) The real bodyguard, the possessed, and the werewolves reveal oneself as the role of the bodyguard on the same day. (iii) Only the real freemasons reveal themselves as the role of the freemasons.

TABLE IV
SETS OF PLAYERS AND ROLES THAT ARE INDEPENDENT OF THE SET OF THE ROLE CANDIDATES $CR(i)$ ON i -TH DAY

ID	expression	description
1	RR	The set of all the roles in the game.
2	$S = \{\text{seer, medium, bodyguard}\}$	The set of all the roles with special skills on the villager side.
3	$Num(r), r \in RR$	Total number of the players with the role r .
4	P	The set of all the players.

reveals oneself as a seer) and that the number of the seer is one as shown in Table I, the role of Liesa is inferred to be the seer and her role is identified.

Fig. 2 also shows a snapshot of the interface of the werewolf game log. On this interface, the date (e.g., the 8th day) and the viewpoint (e.g., the player Dieter's viewpoint) are specified, and then the interface showed role candidates of the players who survive (or who have not survived) on the specified day as well as the utterances and the actions which contribute to inferring roles of those players.

Furthermore, Fig. 3 shows an example of a digest of a werewolf game log, where both the viewpoints of the werewolf side and the villagers side are shown for the final day. On the final day, three players, Joachim, whose role is a werewolf, Dieter, whose role is the possessed, and Jacob, whose role is a villager survive. In this case, if the werewolf is to be voted to be executed, the villager side wins. Otherwise, the werewolf side wins. However, from the viewpoint of the werewolf side, both the werewolf and the possessed have identified each other player's role as the possessed and the werewolf. So, it is very easy for them to vote Jacob, the villager, to be executed. From the viewpoint of the villager side, although Jacob understands that one of the other two players has the role of the werewolf, as shown in the figure, he has not identified which player is actually the werewolf. From Jacob's viewpoint, there still remain two sets of role candidates for the other two players. Then, finally, both Joachim, the werewolf, and Dieter, the possessed, voted Jacob to be executed and the werewolf side won.

V. RULES FOR UPDATING ROLE CANDIDATES OF THE PLAYERS

This section describes the details of the rules for updating the role candidates of each player.

Before describing the details of the rules, Table IV and Table V list the notations of the sets of the players as well as those of the roles we use in this paper. First of all, in this paper, we consider all the possible assignments of the roles to each of the players participating the game. Then, out of those possible assignments of the roles to the players, we pickup one set of the assignments of roles to all the players that is observed on the i -th day, and denote it as $CR(i)$. Here, out of those sets of the players and the roles we used in this paper, some are independent of the set the role candidates

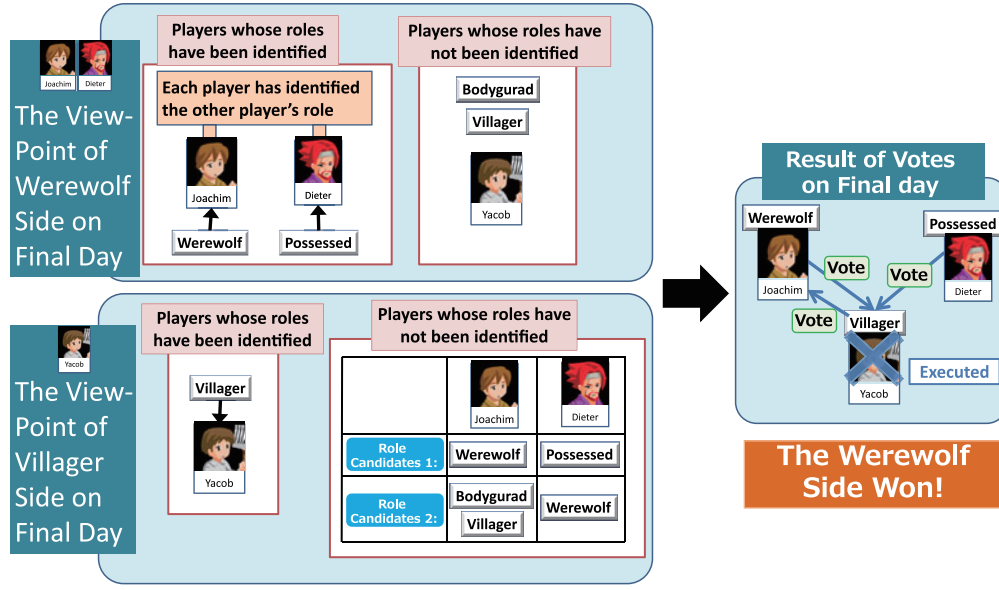


Fig. 3. An Example of A Digest of Werewolf Game Log

TABLE V
SETS OF PLAYERS AND ROLES THAT ARE DEPENDENT ON THE SET OF THE ROLE CANDIDATES $CR(i)$ ON i -TH DAY

ID	expression	description
1	$CR(i) = \{ \langle p, R \rangle \mid p \in P, R \in RR \}$ is the set of the candidate roles for the player p	The set of the role candidates for all the players on the i -th day.
2	$CCR(i) = \{ CR(i), \dots, CR'(i) \}$	The set generated by collecting all the possible combinations $CR(i)$ of the role candidates for all the players.
3	$R(p, i, CR(i))$	The set of the role candidates for the player p , given the set $CR(i)$ of role candidates for all the players on the i -th day.
4	$P(r, CR(i)) = \{ p \in P \mid R(p, i, CR(i)) = \{r\} \}$	The set of players whose role is identified as r , given the set $CR(i)$ of role candidates for all the players on the i -th day.

TABLE VI
RULES FOR UPDATING THE ROLE CANDIDATES OF THE PLAYERS WHO UTTERED OR ACTED

ID	rule
1	After the player p reveals oneself as a role $s \in S$ with special skills on the villager side on the i-th day, for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i-th day, remove any roles other than s, the possessed, and the werewolf from the set $R(p, i, CR(i))$ of the role candidates for the player p. $R(p, i, CR(i)) = R(p, i, CR(i)) \cap \{s, \text{possessed}, \text{werewolf}\}$
2	After the player p reveals oneself as the role of one of the freemasons on the i-th day, for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i-th day, remove any roles other than the freemasons from the set $R(p, i, CR(i))$ of the role candidates for the player p. $R(p, i, CR(i)) = \{\text{freemasons}\}$
3	On the i-th day, after the player p provides the result of the act of the divination (let s be the seer), or after the player p provides the result of the act of the medium (let s be the medium), for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i-th day, divide $CR(i)$ into $CR'(i)$ and $CR''(i)$, where the role of the player p is identified as the role s in $CR(i)$, while the role s is removed from the role candidates of the player p in $CR'(i)$. $CR'(i) = (CR(i) - \{ \langle p, * \rangle \}) \cup \{ \langle p, \{s\} \rangle \}$ $(\langle p, R \rangle \in CR(i)) \quad CR''(i) = (CR(i) - \{ \langle p, * \rangle \}) \cup \{ \langle p, R - \{s\} \rangle \}$ $CCR(i) = (CCR(i) - \{CR(i)\}) \cup \{CR'(i), CR''(i)\}$

TABLE VII
RULES FOR UPDATING THE ROLE CANDIDATES OF THE PLAYERS TO WHOM THE OTHER PLAYER'S ACTION IS DIRECTED

ID	Rule
1	On the i -th day, after the player p is divined to be a werewolf or is identified as a werewolf as the result of the act of the medium, for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i -th day, remove any roles other than the werewolf from the set $R(p, i, CR(i))$ of the role candidates for the player p . $R(p, i, CR(i)) = R(p, i, CR(i)) \cap \{\text{werewolf}\}$
2	On the i -th day, after the player p is divined to be a human or is identified as a human as the result of the act of the medium, for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i -th day, remove the role of the werewolf from the set $R(p, i, CR(i))$ of the role candidates for the player p . $R(p, i, CR(i)) = R(p, i, CR(i)) \cap (RR - \{\text{werewolf}\})$
3	On the i -th day, after the player p is attacked by the werewolves, for all the set $\forall CR(i) (\in CCR(i))$ of role candidates for all the players on the i -th day, remove the role of the werewolf from the set $R(p, i, CR(i))$ of the role candidates for the player p . $R(p, i, CR(i)) = R(p, i, CR(i)) \cap (RR - \{\text{werewolf}\})$

TABLE VIII
RULES FOR RESOLVING INCONSISTENCIES AFTER UPDATING ROLE CANDIDATES OF THE PLAYERS

ID	Rule
1	On the i -th day, within the set $CR(i) (\in CCR(i))$ of role candidates for all the players, when the number of the players whose roles are identified as the role r equals to the total number of the players of the role r ($ P(r, CR(i)) = Num(r)$), from the set $R(p, i, CR(i))$ of the role candidates for the player p other than those whose roles are identified as r ($\forall p \in (P - P(r, CR(i)))$), remove the role r . $R(p, i, CR(i)) = R(p, i, CR(i)) \cap (RR - \{r\})$
2	On the i -th day, within the set $CR(i) (\in CCR(i))$ of role candidates for all the players, for at least one role $\exists r (\in RR)$, if there exists no player whose role candidates include the role r ($\forall p, R \in CR(i), r \notin R$), then remove $CR(i)$ from $CCR(i)$. $CCR(i) = CCR(i) - \{CR(i)\}$
3	On the i -th day, within the set $CR(i) (\in CCR(i))$ of role candidates for all the players, for at least one player $\exists p (\in P)$, if the player p has no role candidate ($R(p, i, CR(i)) = \emptyset$), then remove $CR(i)$ from $CCR(i)$. $CCR(i) = CCR(i) - \{CR(i)\}$
4	On the i -th day, within the set $CR(i) (\in CCR(i))$ of role candidates for all the players, when the number of the players whose role candidates include the role r equals to the total number of the players of the role r ($ P(r, CR(i)) = Num(r)$), for the player p whose role candidates include the role r ($\forall p$ s. t. $r \in R(p, i, CR(i))$), identify the role of the player p as r . $R(p, i, CR(i)) = \{r\}$
5	Let X be the set of the players who revealed themselves as one of the roles with special skills on the villager side (i.e., the seer, the medium, and the bodyguard). Then, $Y = X $ —(the number of the real seer, the real medium, and the real bodyguard) is the number of the members of the subset of X , who are the werewolf or the possessed. Let also Z be the set of the players who are not members of X but whose roles are identified as the werewolf or the possessed. If $Y + Z $ equals to the total number of the werewolves and the possessed, then, from the role candidates of the players who are not members of $X \cup Y$, remove the werewolf and the possessed.
6	For the players who do not revealed themselves as a seer on or before the third day, remove the role of the seer from their role candidates when the fourth day starts. For the players who do not revealed themselves as a medium on or before the third day, remove the role of the medium from their role candidates when the fourth day starts. After a player p revealed oneself as a bodyguard on the i -th day, remove the role of the bodyguard from the role candidates of the players other than p on the $i+1$ -th day. After two players p and q revealed themselves as freemasons, remove the role of the freemasons from the role candidates of the players other than p and q .
7	Let X be the set of the players who survive and are identified as the werewolf. If $ X $ equals to the maximum number of the werewolves before the werewolf side wins (i.e., $Y - 1$, where Y be the minimum integer that is greater than or equal to the half of the number of the humans who survive), then, remove the role of the werewolf from the role candidates of the players who survive and are not members of X .

$CR(i)$ for all the players on the i -th day, which are listed in Table IV. Others are, on the other hand, quite dependent on the set $CR(i)$, which are listed in Table V.

Next, Table VI, Table VII, and Table VIII list all the rules for updating the role candidates for the players. Table VI considers the cases where the player reveals oneself as having a role with special skills on the villager side such as the seer, the medium, and the bodyguard. Those utterances have quite strong influences on the role candidates of those who reveal oneself. Table VII, on the other hand, considers the cases where updating the role candidates of the players to whom the other player's action such as divination, medium, and attack is directed. The acts of divination and medium

provide information on the role of the selected players, while the act of attack by werewolves provide information that the roles of those attacked are humans. Finally, after updating the role candidates of the players by the rules listed in Table VI and Table VII, it is usually the case that inconsistencies of the role candidates of the players may arise. In such cases, rules listed in Table VIII are used to resolve those inconsistencies.

VI. UPDATING ROLE CANDIDATES OF THE PLAYERS

A. The Procedure

This section describes the procedure of applying the rules for updating the role candidates of the players presented in the previous section.

1) *Initialization*: First, let i be 1 and we describe how we initialize the set $CR(i)$ of role candidates for all the players in the beginning of the first day.

The following equation for $CR_0(i = 1)$ denotes the initial definition of $CR(i)$, where, for all the players $p \in P$, the role candidates are defined as the set RR of all the roles in the game.

$$CR_0(i = 1) = \{ \langle p_1, RR \rangle, \dots, \langle p_n, RR \rangle \} \quad (n = |P|)$$

Henceforth, with this initial definition $CR_0(i = 1)$, the following equation also holds.

$$\forall p \in P, R(p, i = 1, CR_0(i = 1)) = RR$$

Also, the set $CCR(i)$ generated by collecting all the possible combination $CR(i)$ of the role candidates for all the players is initialized as the following equation.

$$CCR(i = 1) = \{ CR_0(i = 1) \}$$

2) *Updates on i -th Day*: The following procedure describes how to update the role candidates of all the players on i -th day.

- (i) Repeat the following procedure of processing all the utterances U and the actions A one by one on i -th day.
 - a) If the utterance U or the action A is among revealing oneself, divination, medium, or attack, apply an appropriate rule of Table VI or Table VII.
 - b) Repeat applying any rule of Table VIII while role candidates of at least one player do not converge.
- (ii) If the i -th day is the final day of the game, stop the procedure. Otherwise, update the date of the game by simply updating the set $CCR(i)$ according to the following formula.

$$CCR(i + 1) = CCR(i)$$

Then, continue the whole procedure on $i+1$ -th day.

B. An Example

This section describes the way each player's role is inferred in Fig. 1 as presented in the previous section according to the procedure of updating the role candidates of the players.

First, the role candidates for Liesa is initialized as the set RR of all the roles in the game. Next, after Liesa revealed herself as the role of a seer, the rule with ID=1 in Table VI is applied and the role candidates for Liesa are to be {seer, possessed, werewolf}. Suppose that no player other than Liesa reveals oneself as the role of the seer on or before the third day. Then, the rule with ID=6 in Table VIII is applied and from the role candidates of all the other players, the role of the seer is removed.

Finally, the requirement of applying the rule with ID=4 in Table VIII is satisfied since, for players other than Liesa, the set of the role candidates do not include the role of the seer, and the total number of the players with the role of the seer

$Num(seer)$ is 1. Then, the rule with ID=4 in Table VIII is applied and the role of Liesa is identified as the seer.

As for the evaluation of the proposed framework of inferring the role candidates of the players, we apply the proposed framework to one whole game log which consist of around 3,000 utterances, where we confirm that the proposed set of rules is quite effective in inferring all the participating players' role candidates.

VII. RELATED WORK

As the related work on developing a computer agent program which participates in the werewolf game, most studies have examined the face-to-face werewolf game and analyzed non-verbal audio cues, physical gestures, and conversational features such as speaker turns. Chittaranjan and Hung [3] analyzed the audio-visual recordings of the face-to-face werewolf game and trained a classifier for detecting lying or suspicious behavior using non-verbal audio cues in a social context and for predicting players' decisions by analyzing speaker turns. Katagami et al. [4], [5] also studied the effect of non-verbal information in the face-to-face werewolf game. Kobayashi et al. [6] developed a werewolf match system for human players which is mediated with lifelike agents and has both functions of conversation by chat communication and non-verbal communication.

VIII. CONCLUSION

The underlying goal of this paper is to construct an agent that can analyze the participating players' utterances and play the werewolf game as if it is a human. For the first step of this underlying goal, given a specific player participating in the wolf game, this paper studied how to generate a digest of inferring other players' roles from the viewpoint of the given specific player. We developed a set of inference rules and applied them to infer the participating players' roles from a real werewolf game log. Future work includes much larger scale evaluation of the proposed framework of inferring the players' role candidates by applying the proposed framework to much larger number of game logs.

REFERENCES

- [1] T. Shinoda, F. Chokai, D. Katagami, H. Osawa, and T. Inaba, "“Are you a Werewolf?” becomes a standard problem for general artificial intelligence," in *Proc. 28th Annual Conf. JSAI*, 2014, (in Japanese).
- [2] F. Toriumi, K. Kajiware, H. Osawa, T. Inaba, D. Katagami, and T. Shinoda, "Development of AI wolf server," *Proc. 19th Game Programming Workshop*, pp. 127–132, 2014, (in Japanese).
- [3] G. Chittaranjan and H. Hung, "Are you a werewolf? detecting deceptive roles and outcomes in a conversational role-playing game," in *Proc. ICASSP*, 2010, pp. 5334–5337.
- [4] D. Katagami, S. Takaku, M. Inaba, H. Osawa, K. Shinoda, J. Nishino, and F. Toriumi, "Investigation of the effects of nonverbal information on werewolf," in *Proc. FUZZ-IEEE*, 2014, pp. 982–987.
- [5] D. Katagami, M. Kanazawa, F. Toriumi, H. Osawa, M. Inaba, and K. Shinoda, "Movement design of a life-like agent for the werewolf game," in *Proc. FUZZ-IEEE*, 2015, pp. 1–7.
- [6] Y. Kobayashi, H. Osawa, M. Inaba, K. Shinoda, F. Toriumi, and D. Katagami, "Development of werewolf match system for human players mediated with lifelike agents," in *Proc. 2nd HAI*, 2014, pp. 205–207.