

各プレイヤー視点での役職確定情報に基づく 人狼ログ・ダイジェストの作成*

林友超[†] 馬場 瑞穂[‡] 宇津呂 武仁[‡]

筑波大学大学院 システム情報工学研究科[†] 筑波大学システム情報系[‡]

1 はじめに

「昼間に人間のふりをして、夜に正体を現して、人間を襲撃する人狼が村に現れた。村人達はお互いを疑いつつ、毎日投票で一番疑わしい一人を処刑する。果たして村人は全ての人狼を排除できるか?」

これが人狼ゲームの簡単なストーリーである。近年、人狼ゲームは日本国内でもよく知られている。例えば、舞台劇中でアドリブで人狼ゲームをプレーする人狼 TLPT¹や、テレビ番組で芸能人を集めて、人狼ゲームをプレーするもの²等が知られている。

一方、人工知能の研究分野においては、2014年に、人狼ゲームは汎用人工知能の標準問題の一つとあるという提案がなされた [4]。また、2015年8月には、人狼プロジェクト³が公開したサーバ [5]を用いて、第一回の人狼知能大会が開催された⁴。このように、人狼ゲームを題材とする人工知能研究が近年行われている。しかし、現時点での人狼ゲームを題材とする人工知能研究の多くでは、プレイヤー間における自然言語会話の内容を理解すること、その会話内容の理解をふまえてプレイヤーの役職を推論すること、高度な推論を介して襲撃・処刑対象のプレイヤーを決定すること等、人間の推論レベルを目指した人狼 AI の研究のレベルにまでは到達していない。

この現状をふまえて、本論文では、人間のレベルの人狼 AI を実現するための基礎的な要素技術の一つとして、人間のプレイヤーによる人狼ゲームのプレーログを自然言語解析し、各プレイヤーの視点において、他プレイヤーの役職が確定していく様子を人狼ログ・

ダイジェストとして作成する方式を提案する。この方式においては、各プレイヤーによってなされる多様な役職宣言発言を同定し、この発言内容と人狼ゲームのルール、および、人狼ゲームにおける前提知識として用いられる常識に基づく推論により、他プレイヤーの役職の可能性を絞り込む過程をダイジェストとして可視化する。本論文で利用する役職宣言発言の典型例としては、本物の占い師による自らの役職宣言発言 (占い師 CO)、あるいは、偽りの役職宣言発言によって人狼ゲームを有利に進めようとする人狼・狂人によってなされる偽りの占い師役職宣言発言 (占い師 CO) が挙げられる。今後、人狼 AI を実現する過程においても、本論文で提案する方式と同等レベルの推論に基づいて、他プレイヤーの役職を確定する技術を実現することは不可避であると考えられる。したがって、本論文の成果は、今後の人狼 AI 研究においても、重要な要素技術の一つとして位置付けることができる。

2 人狼

人狼ゲームにおいては、ゲーム開始時に、各プレイヤーが一つの役職を与えられて、役職に応じて「人狼陣営」と「村人陣営」に分けられ、各自の陣営の勝利を目指してプレーする。人狼陣営は、自身の勝利のために、事実と矛盾する発言を行い、村人陣営側を混乱させる。それに対して、村人陣営は、議論の内容の真偽を判別して、人狼の正体を見破り、議論を通じて人狼を処刑することを目指す。

人狼ゲームには、何通りかの人数制が設けられているが、一例として、プレイヤー数が16人の場合の役職一覧を表1に示す。表中、「狂人」は、占い師や霊能者による判定結果は「人間」となり、ゲームの勝敗判定においても「人間」として扱われるが、「狂人」自身にとっては、人狼陣営の勝利が「狂人」自身の勝利となるという特別な役職である。

その他、人狼ゲームにおけるルールおよび常識の一覧を表2に示す。表中、ゲームのルールとは、決して覆ることはなく、各プレイヤーが確実にそれに従って

*Generating a Werewolf Game Log Digest based on Crucial Utterances about Other Players' Roles in Each Player's View

[†]Youchao Lin, Mizuho Baba, Graduate School of Systems and Information Engineering, University of Tsukuba

[‡]Takehito Utsuro, Faculty of Engineering, Information and Systems, University of Tsukuba

¹<http://7th-castle.com/jinrou/index.php>

²<http://www.fujitv.co.jp/jinroh/index.html>

³<http://www.aiwolf.org/>

⁴<http://cedec.cesa.or.jp/2015/session/AC/7649.html>

表 1: 人狼ゲームにおける役職の一覧 (プレイヤー数 16 人の場合)

陣営	生存人数カウントでの扱い	役職	説明	人数
村人	人間	村人	特別な役職を持たない人間である.	7
		占い師	村人陣営の人間である, 毎夜, 一人を占って, その人が人狼なのか, 人間なのかを知り, 公表することができる.	1
		霊能者	村人陣営の人間である. 投票により処刑されたプレイヤーが, 人狼なのか人間なのかを知り, 公表することができる.	1
		狩人	村人陣営の人間である. 毎夜, 一人だけを人狼の襲撃から守ることができる. 護衛対象が実際に襲撃をうけたかどうかを知ることはできない.	1
		共有者	村人陣営の人間である. もう一人の共有者が誰かを知ることができる.	2
人狼	人狼	狂人	人狼陣営の人間である. 人狼の勝利が狂人の勝利となる. 狂人と人狼はお互いに正体を知らない.	1
		人狼	毎夜, 一人だけ人間を襲撃できる. 人狼同士にしか聞こえない会話が可能である.	3
計				16

表 2: 人狼ゲームのルールおよび常識

(a) ルール

1	占い師, 霊能者, 狩人, 狂人はそれぞれ一人だけ.
2	共有者は二人でかつ相手の正体を互いに知っている.
3	人狼同士は互いの正体を知っている.
4	人狼は人狼を襲撃することができない.
5	人狼の人数が人間の人数以上になると, 人狼陣営の勝利.
6	人狼を全て処刑できれば, 村人陣営の勝利.

(b) 常識

1	村人, 占い師, 霊能者, 狩人, 共有者は村の勝利のために, 事実と矛盾しない情報を発言する.
2	占い師, 霊能者は村のために名乗り出る.
3	人狼, 狂人は人狼陣営の勝利のために事実と矛盾する情報を発言する.

ゲームを進行する規則を指す. 一方, ゲームの常識とは, あらかじめ明示的に規定されているわけではないが, 一般的には, 各プレイヤーにとって周知の知識であり, 通常はその知識の範囲内で行動すると考えているものである. 例えば, 「1. 村人陣営のプレイヤーは事実と矛盾しない情報を発言する」の常識について, 厳密には, 村人陣営の村人も事実と矛盾する情報を発言する可能性もあり得るが, 事実と矛盾する情報を発言することによって, 村人陣営の他の人間の混乱を招き, 結果的に村人陣営の不利益となる可能性が高い. したがって, 本論文においては, これらの一連の常識についても, 役職推論における前提知識として用いる.

3 人狼ログ

本論文では, ネット上の人狼ゲームサイトである人狼BBS⁵において公開されているゲームログを対象として, 人狼ログ・ダイジェストを作成する. 人狼BBS

サイトにおいては, 各プレイヤーは匿名で人狼ゲームをプレーする. 各プレイヤーは専用の掲示板に文字を入力して, 議論を行う. 各プレイヤーは表示される発言ログに対して真偽判断を行いながら, ゲームを進行する.

4 役職特定のための情報

本論文では, 表 3 に示す計 8 種類の役職宣言発言を情報源として役職推論を行う⁶. 各発言に加えて, 表 2 に示したゲームのルールおよび常識のいくつかを適用することにより, 表 3 の「得られる情報」が導出される.

5 各プレイヤー視点での人狼ログ・ダイジェスト

5.1 基本的な考え方

本論文の人狼ログ・ダイジェスト作成における役職推論の基本的な考え方は, プレイヤーの多様な役職宣言発言に加えて, ゲームのルールおよび常識のいくつかを適用することにより, 役職推論を行い, 各プレイヤーの役職の可能性を絞り込むというものである. 図 1 に一例を示す. 図 1 では, ある時点で, リーザだけが占い師 CO し, 占い師であると名乗り出ており, 一方, 他のどのプレイヤーも, 占い師であるとは名乗り出ていない. この時点での, 「リーザだけが占い師と名乗り出ている」という情報と, ゲームの常識「占い師は, 村人陣営の勝利のために名乗り出る」, および, ゲームのルール「占い師は一人だけ」の 2 つの知識に基づいて推論を行った結果, 占い師候補はリーザ一人だけとなり, リーザが本物の占い師であると確定する.

⁶役職宣言発言の典型例としては, 役職 CO 宣言が挙げられる. CO とは "Coming Out" の略で, 「名乗り出る」という意味.

⁵<http://ninjin002.x0.com/wolff/>

表 3: 役職宣言発言類一覧

種類	パターン	例	得られる情報
占い師 CO (占い師と名乗り出る)	私は占い師です.	リザちゃんが占い師ですよ.	COした人は占い師, 人狼, 狂人のうちいずれか. COしなかった人は占い師ではない.
占い結果	[名前] は [人間/人狼] です.	ヨア兄は人狼だったですう	占われた人は [人間/人狼].
霊能者 CO (霊能者と名乗り出る)	私は霊能者です.	俺が霊能者だ	COした人は霊能者, 人狼, 狂人のうちいずれか. COしなかった人は霊能者ではない.
霊能判定結果	[名前] は [人間/人狼] です.	リーザは人間だったぜ	霊能判定結果の対象は [人間/人狼].
狩人 CO (狩人と名乗り出る)	私は狩人です.	オレが狩人だからだ	COした人は狩人, 人狼, 狂人のうちいずれか.
狩人護衛対象	[名前] を守りました.	護衛したのはヨアヒム	襲撃で死亡した人がいなかった日の護衛対象が人間である.
共有者 CO (共有者と名乗り出る)	私は共有者です.	私は【共有者】です.	COした人は共有者, 人狼, 狂人のうちいずれか.
共有者相手	私の相手は [名前] です.	そう, あたしがジムソンの相方共有者さ.	[名前] は共有者である.

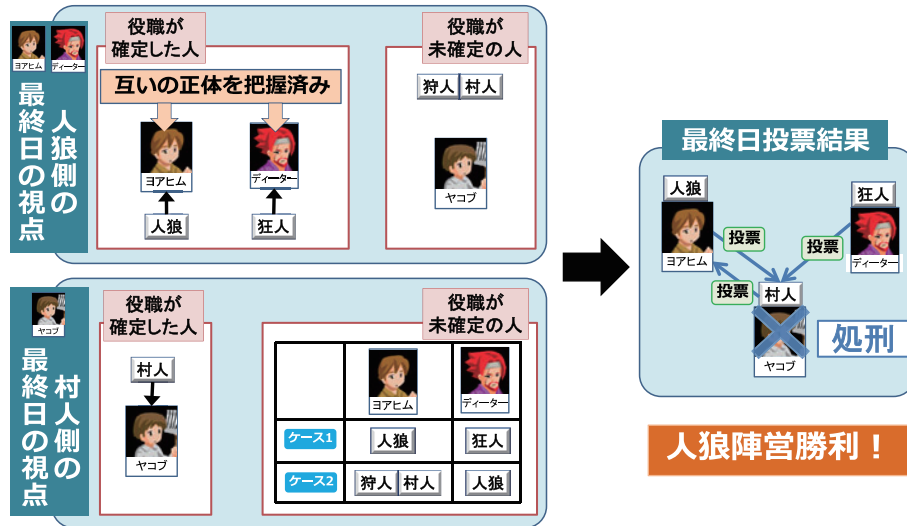


図 2: 人狼ログのダイジェストの例

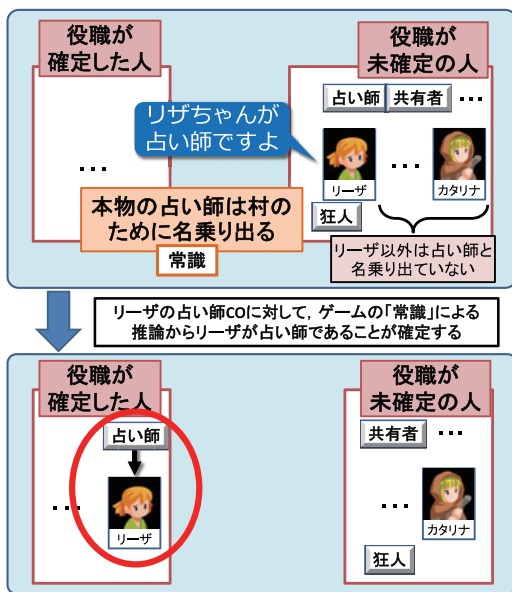


図 1: 役職推論の例

5.2 具体例

本節では, 人狼 BBS から実際に収集可能な F539 村のログを対象としてダイジェストを作成する過程を示す. この村のプレイヤー数は 16 人で, 役職構成は表 1 の通りである. 村の最終日に生き残ったのは, ヨアヒム (人狼), ディーター (狂人), およびヤコブ (村人) の三人である. 最終日開始時点で, この三人の視点からの他プレイヤーの役職確定状況は, 図 2 のダイジェスト例に示す通りである. 以下では, 各プレイヤーの視点において, 他プレイヤーの役職が確定する過程を順を追って示す.

5.2.1 ヨアヒム (人狼) 視点

一日目 ディーターとリーザが占い師 CO して名乗り出た. 人狼であるヨアヒムには, どちらも人間であると分かっており, 表 2(a) のルール 1, 表 2(b) の常識 2 および常識 3 より, どちらかが本物の占

い師でどちらかが狂人であると分かった。

三日目 二日目の占い結果が三日目の朝発表された段階で、ディーターが「ヨアヒムは人間」と発言し、逆に、リーザが「ヨアヒムは人狼」と発言した。表 2(b) の常識 1 より、事実と矛盾する「ヨアヒムは人間」を占い結果として発言したディーターが狂人であると分かった。

5.2.2 ディーター (狂人) 視点

一日目 自分以外に、ヴァルターとリーザが占い師 CO して名乗り出た。狂人であるディーターには、表 2(a) のルール 1、表 2(b) の常識 2 および常識 3 より、どちらかが本物の占い師で残りのもう一人が人狼であると分かった。

三日目 二日目の占い結果が三日目の朝発表された段階で、リーザが「ヨアヒムは人狼」と発言した。

四日目 三日目の処刑対象としてリーザが処刑された。これに対し、二日目に霊能者 CO して名乗り出たトーマスとオットーの二人とも霊能判定結果として「リーザは人間」と発言した。表 2(b) の常識 1 および常識 2 より、トーマスとオットーのどちらかは本物の霊能者であり、その霊能判定結果である「リーザは人間」は事実であると分かった。このことと、一日目の推論結果から、リーザが本物の占い師であると分かった。

このことと、表 2(b) の常識 1、および、三日目朝の占い結果「ヨアヒムは人狼」から、「ヨアヒムは人狼」は事実であると分かった。

5.2.3 ヤコブ (村人) 視点

村人陣営のヤコブの視点では、最終日の段階でも、ヨアヒムとディーターのどちらかが人狼で、残りのもう一人が人間であることまでしか分からなかった。

5.2.4 ゲームの勝敗

以上の結果から、最終日開始段階で、人狼陣営の人狼ヨアヒムと狂人ディーターに対しては、お互いの正体が判明していたため、最終日の処刑対象への投票時に、票を合わせて村人陣営の村人ヤコブに投票すればゲームに勝利するという結論が得られた⁷。以上のように、本論文の人狼ログ・ダイジェストの枠組みによって得られる情報に基づき、プレイヤーの投票行動が容易に予測可能であることが示された。

⁷実際に、この村においては、最終的に、ヨアヒムとディーターが村人陣営のヤコブに投票して、勝利した。

6 関連研究

本論文に関連して、文献 [3] においては、人狼ゲームの性質の一つとして、「プレイヤーは自身の役職を客観的に証明することができない」と述べている。これに対して、本論文では、人狼ゲームにおけるルールおよび常識をいくつか設定することにより、人狼ゲーム・ログ中に観測される役職確定情報を手がかりとして他プレイヤーの役職を推論することが可能であることを示した。

また、文献 [2] においては、村人陣営、人狼陣営の各々において、他人の意見に対する同調・反駁の確率分布を測定し、村人陣営・人狼陣営を識別する手がかりとして有効であることを示した。文献 [1] においては、人狼 BBS ログ中の役職宣言発言を抽出し、人狼エージェントにおいて襲撃対象・処刑対象を選定する確率モデルのパラメータとして組み込む方式を提案した。この研究では、人狼エージェントにおける行動選択確率モデルのパラメータ調整を目的としているが、一方、本論文では、各プレイヤー視点での役職推論の過程をダイジェストとして可視化することを目的としている。

7 おわりに

本論文では、人狼ゲームにおける AI 実現の一環として、人狼ゲームのルールおよび常識を推論規則として用い、人狼ログ中の役職宣言発言を手がかりとして役職推論を行う過程を人狼ログ・ダイジェストとして可視化する方式を提案した。今後の課題として、人狼ログ中の役職宣言発言を自動抽出して、実際に役職推論を行う過程を実装することが挙げられる。

謝辞

本研究を行うにあたり、人狼 BBS のデータ使用を許可していただいた ninjin 氏に感謝いたします。

参考文献

- [1] 平田佑也, 稲葉通将, 高橋健一. プレイログから獲得した行動選択確率を用いた人狼ゲームのシミュレーション. 第 29 回人工知能学会全国大会論文集, 2015.
- [2] 稲葉通将, 鳥海不二夫, 大澤博隆, 片上大輔, 篠田孝祐, 西野順二. 同調と反駁に着目した人狼ゲームの分析. 第 28 回人工知能学会全国大会論文集, 2014.
- [3] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 松原仁, 小林潤平, 関口隆, 新堀英二, 川嶋稔夫. 人狼知能プロジェクト. 人工知能学会誌, Vol. 30, No. 1, pp. 65–73, 2015.
- [4] 篠田孝祐, 鳥海不二夫, 片上大輔, 大澤博隆, 稲葉通将. 汎用人工知能の標準問題としての人狼ゲーム. 第 28 回人工知能学会全国大会論文集, 2014.
- [5] 鳥海不二夫, 梶原健吾, 大澤博隆, 稲葉通将, 片上大輔, 篠田孝祐. 人狼知能サーバの構築. ゲームプログラミングワークショップ 2014 論文集, pp. 127–132, 2014.