

人狼ゲームをプレイする自律エージェントの戦略に関する検討

Study of strategies for autonomous agents that play "werewolf game"

永山 翔滋^{1*} 阿部 穰太郎¹ 大矢 康介¹
Shoji Nagayama¹ Jotaro Abe¹ Kosuke Oya¹
阪本 浩太郎^{1,3} 渋谷 英潔² 森 辰則² 神門 典子^{3,4}
Kotaro Sakamoto^{1,3} Hideyuki Sibuki² Tatunori Mori² Noriko Kando^{3,4}

¹ 横浜国立大学 大学院 環境情報学府 ² 同 環境情報研究院

¹Yokohama National University ²*ditto*

³ 国立情報学研究所 ⁴ 総合研究大学院大学

³National Institute of Informatics ⁴SOKENDAI

Abstract: "Werewolf game" is a popular multiplayer game in which "villagers" are trying to figure out who is a "werewolf" through conversation. Werewolves are usually pretending to be villagers. In this paper, we studied conversation in game logs in order to investigate how werewolves' cooperation contributed to winning percentage of werewolves' team. Since the number of "whispers" that are utterances via werewolves' private chat may be regarded as a measure of the werewolves' cooperation, we investigated the relation between the number of whispers and the winning percentage. As the result, we observed that the winning percentage of werewolves' team increased by 63 percentage point at most when the number of whispers of at least two werewolves was more than 106.

1 はじめに

人狼ゲームとは、対話を通して「村人」の中に潜伏した「人狼」を見つけ出す対戦型の多人数ゲームであり、近年、人狼知能プロジェクト¹など研究テーマとしても注目されている。人狼知能プロジェクトでは、「人間と自然なコミュニケーションを取りながら人狼をプレイするエージェント(AI)の構築」を究極の目標としており、エージェント同士の対戦成績を一つの指標として人工知能の研究開発に取り組んでいる。

人狼ゲームの詳細は3節に後述するが、人狼ゲームは村人陣営と人狼陣営に分かれて戦うチーム戦であり、対話というインタラクティブなコミュニケーションを通してチームの勝利条件を達成しようとする。しかしながら、ゲーム中の対話は基本的に全てのプレイヤーに伝わるオープンなものであり、特定のプレイヤーとだけ対話することはできない。そのため、味方陣営だけで相談するといったチーム全体での連携は困難であり、個人ごとの思惑が錯綜する点が人狼ゲームの特徴である。一方で、「人狼」となったプレイヤーには、「囁き」と呼ばれる人狼同士だけで対話ができる特別な対話チャンネルが用意されており、人狼同士の連携を可能にできる。これは、人狼以外のプレイヤーにはない強力なアドバンテージであり、人狼陣営が勝利するためには重要な要素であると考えられる。人狼ゲームの対話を分析した研究として、文献[1, 2]などが挙げられる

が、人狼同士の連携に焦点をあてたものは存在しない。我々は、文献[3]において、最低一人の人狼の発話数が101を超えることで人狼陣営が勝利する可能性が最大65ポイントに上がるという知見を得ているが、発話数を調整するためにどのように人狼同士が連携するかについては考察できていなかった。以上の背景から、本稿では「人狼」エージェントに焦点をあて、人狼同士の囁きが勝敗にどのような影響を与えるかを分析する。

本稿の構成は以下の通りである。2節で関連研究を述べ、本稿の位置付けを明確にする。3節で人狼ゲームの役職とルールを説明した後、人狼プレイヤーの基本戦術であるステルス人狼の観点から発話と注目が勝敗に影響を与える要因であることを述べる。4節では、分析対象とした人狼BBSについて説明をする。5節で文献[3]の分析結果を紹介した後、6節で、人狼同士の囁きが勝敗とどのような関係があるかを考察する。7節は結論である。

2 関連研究

人狼ゲームの対話を分析した研究として、稲葉らの研究[1, 2]がある。稲葉らの研究では、分析対象として、オンラインで行う掲示板型人狼ゲームである人狼BBS²のログを用いており、発話内容にその種類を分類するタグを付与し集計することで、タグの頻度により、ゲームの勝敗、襲撃や処刑される人の特徴を分析している。稲葉らは発話の内容に着目しており、一方で対象にはほとんど着目していない。対象とは、発話した人

*連絡先：横浜国立大学環境情報学府
〒240-8501 横浜市保土ヶ谷区常盤台 79-7 森 辰則研究室
E-mail: nagayama@forest.eis.ynu.ac.jp

¹<http://aiwolf.org/>

²<http://www.wolfg.x0.com/>

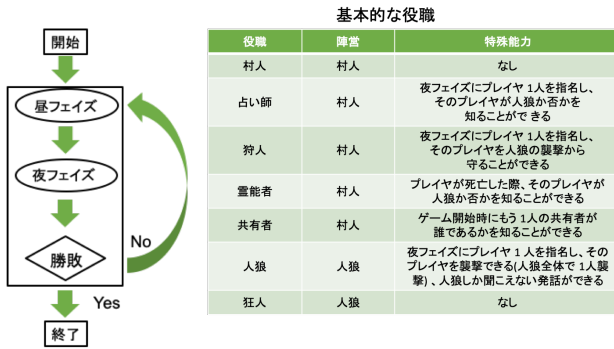


図 1: ゲームの流れと主な役職

が誰であるか、また誰に向けて発話しているかということである。例えば、発話内容タグには占い師のことについて語るタグ (fortune telling) があるが、占い師が本物か偽物かを区別しない。占い師を信用する発話があった場合、本物の占い師であれば村人陣営の勝利に傾いていることを意味するが、偽物であれば人狼陣営の勝利に傾いている。そのため、我々は文献 [3] において、対象を明確にした上で勝敗に基づく分析を行った。その結果、発話数においては人狼の発話数が 101 を超え、また注目数においては占い師の注目数が 168 を下回る場合、人狼の勝率が高いことがわかった。本稿では、文献 [3] での分析をさらに発展させ、人狼の連携度についての分析を行う。

3 人狼ゲーム

3.1 役職とルール

図 1 にゲームの流れと基本的な役職³を示す。役職は、ゲーム開始時に各プレイヤーにランダムに割り当てられ、他のプレイヤーからは分らないようになっている。人狼ゲームはチーム戦で村人陣営と人狼陣営に分かれて戦う。人狼以外のプレイヤー数が人狼プレイヤーの数以下になると人狼陣営の勝利となり、それまでに人狼プレイヤーが全て処刑されると村人陣営の勝利となる。幾つかの役職には図に示した特殊能力があり、これを手がかりとして人狼プレイヤーを探す (村人陣営)、または見つからないようにする (人狼陣営)、というのがゲームの基本である。しかしながら、殆どのプレイヤーは特殊能力を持たない村人であるため、どちらの陣営にとっても彼らを如何にして説得するかが勝敗の鍵となる。

ゲームは「昼」と「夜」のフェイズに分かれて進行する。昼フェイズでは、誰が人狼であるかを議論し、人狼と思われるプレイヤー1人を多数決で処刑する (ゲームから脱落させる)。夜フェイズでは、自分の陣営を優位にするために特殊能力を行使する。例えば、占い師であれば任意のプレイヤー1人の役職が人狼であるか否かを知ることができ、人狼プレイヤーであれば任意のプレイヤー1人を襲撃する (ゲームから脱落させる) ことができる。以上で「1日」が終わり、勝敗がつ

³人狼ゲームには多くの特殊ルールがあり、ここで挙げた役職以外にも存在する。



図 2: 人狼 BBS の例

かなければ「翌日」となる。つまり、1日に2人 (昼と夜で1人ずつ) プレイヤーが減っていくことになる⁴。

3.2 人狼の基本戦術

3.1 節のルールを踏まえて、人狼プレイヤーの視点からは、以下の二種類の基本戦術が考えられる。占い師や霊媒師などの役職を騙ることで特殊能力による (と思わせた) 誤情報を流す戦術 (騙り人狼) と、特殊能力がない村人として振る舞い、目立つ行動をしないことで処刑や占いなどを逃れる戦術 (ステルス人狼) である。騙り人狼は、村人を誤った結論に能動的に誘導できる利点があるが、騙られた相手からは人狼陣営⁵であることが分かってしまい、処刑や占いの対象になりやすいという欠点もある。一方で、ステルス人狼は、議論の誘導という点では受動的であるが、処刑や占いの対象に直接働きかけない⁶ため、相手に人狼であることが疑われにくい。ゲーム中に複数の人狼がいる場合、上記の戦術は分担されることが多い。全員が騙り人狼を行うと、本来1人しかいないはずの占い師などの役職が複数人存在することになり、役職を持っている全員を処刑することで人狼陣営が敗北する。また、全員がステルス人狼を行うと、占い師や霊媒師が本物でほぼ確定⁷し、村人陣営が有利となる。したがって、複数の人狼がいる場合、両方の戦術を分担する戦略がとられやすい。

⁴狩人が襲撃から守るなどの例外は存在するが基本的には2人ずつ減っていく。

⁵狂人の可能性もあるため人狼であることまでは分らない。

⁶人狼が占い師として騙りに出ている場合、処刑や占いの対象を直接左右する発言を行うことができる。

⁷狂人が騙る可能性がある。

表 1: 各村のデータ数と役職の人数

キャラクター数	ファイル数	役職ごとのキャラクター数						
		村人	占い師	狩人	霊能者	人狼	狂人	NPC
14 人	89	6 人	1 人	1 人	1 人	3 人	1 人	1 人
15 人	33	7 人	1 人	1 人	1 人	3 人	1 人	1 人
16 人	309	8 人	1 人	1 人	1 人	3 人	1 人	1 人
全体	431							

本稿では、プレイヤー個人の戦術に着目し、ステルス人狼に焦点をあてる。ステルス人狼の「村人として振る舞い、目立つ行動をしない」という基本行動は、実際のゲームにおいて簡単ではない。村人は特殊能力をもたないが、これは昼フェイズの議論において黙っていることを意味しない。ゲームの序盤では、これといった根拠がない状況で誰かを処刑しなければならない場面がしばしばあり、そういった状況では村人陣営にとってあまり有益な情報をもたらさないプレイヤーから処刑することが多い。結果、議論を黙って聞いているだけだったり、無益な発言をするだけのプレイヤーは処刑の対象となりやすい。したがって、ステルス人狼は、ある程度議論に参加して、村人陣営に有益である（と思わせる）発言をしなくてはならない。

ステルス人狼が、さらに難しいのは、有益すぎる発言もできないということである。これは本当の人狼プレイヤーを特定させる発言ができないというだけではない。村人陣営にとって非常に有益なプレイヤーであると思われた場合、本当に村人陣営であれば人狼から襲撃されるはずである。それにも関わらず、襲撃されていないという事実は、本当は人狼陣営であるという疑いを抱かせることになり、処刑や占いの対象となりやすくなる。したがって、ステルス人狼を実現するためには、「村人陣営に有益な発言をしつつも目立たない」というバランスを保った発言をする必要がある。

4 人狼 BBS

インターネットの掲示板を利用した人狼ゲームとして人狼 BBS がある。図 2 に人狼 BBS の画面を示す。1 つの村⁸にはノンプレイヤーキャラクター (NPC) 1 人を含めた 10 人から 16 人⁹が参加し、ゲーム内の 1 日を現実の 1 日で行う。各プレイヤーが 1 日にできる最大発言数は 20 回であり、またテキストベースで行うためノンバーバルコミュニケーション情報が一切考慮されないのも特徴である。人狼 BBS において、人狼は専用のチャット (囁き) で内通ができる。図 2 中の赤い吹き出しが囁きにあたり、人狼であるエルザとリーザが戦略について相談している。ゲームの早い段階で、人狼たちは囁きを使って、採用する戦略や、騙り人狼とステルス人狼の分担などを相談することが一般的である。

本稿では、分析対象として人狼 BBS:G 村から python の beautifulsoup¹⁰を使ってログを収集した。人狼が 3 人存在する村¹¹、かつ、ゲーム中に処刑や襲撃以外でプレイヤーが脱落していない¹²村を対象として、1 つの

⁸人狼 BBS では 1 回のゲームを 1 つの村と呼称する。

⁹実際のプレイヤー数は 9 人から 15 人となる。

¹⁰<https://github.com/waylan/beautifulsoup>

¹¹キャラクター数が 13 人から 16 人の村が該当する。

¹²1 日に 1 回も発言しなかったプレイヤーは強制的に脱落させられる。また、自らゲームを止める場合もある。

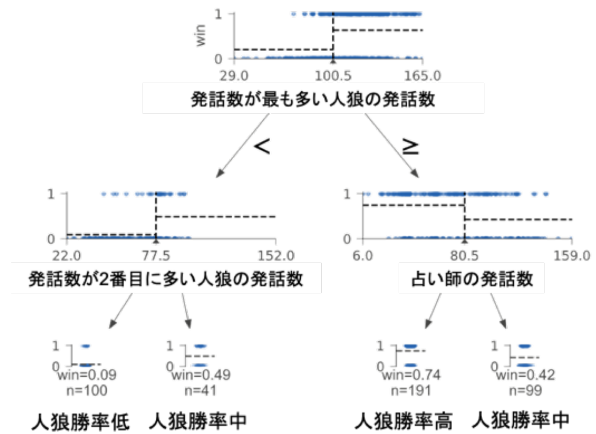


図 3: 発言数に基づく決定木 (ゲーム全体)

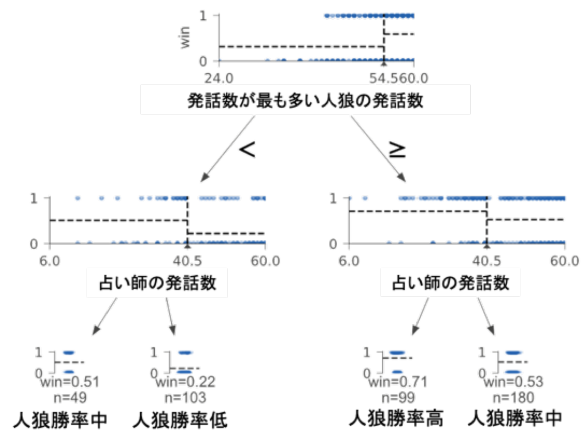


図 4: 発言数に基づく決定木 (ゲーム前半)

村を 1 ファイルとして収集した。収集条件を満たしたのは、表 1 の通り、14 人から 16 人の村について、全体で 431 ファイル (243MB) となった。各ファイルにはゲーム開始前のプロローグ部分が含まれているが、役職が割り当てられる前の雑談であるため、本稿では分析対象からプロローグ部分を除外した。除外後の 1 ファイルあたりの平均発言数は 70.7 発言となった。

5 プレイヤの発言数/注目数と勝率との関係の分析

本節では、文献 [3] の結果を紹介することで、発言数や注目数と勝率との関係を述べる。3.2 節で述べたように、ステルス人狼は「村人陣営に有益な発言をしつつ

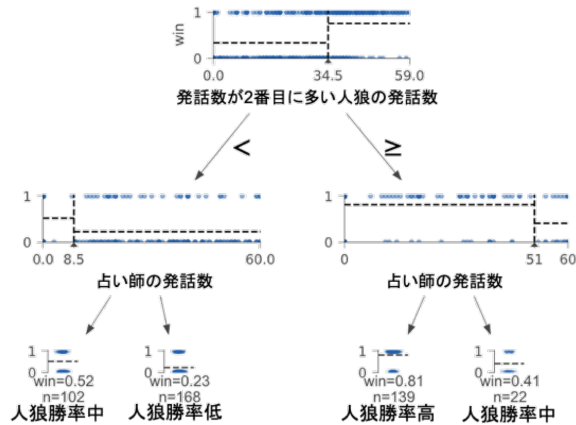


図 5: 発言数に基づく決定木 (ゲーム後半)

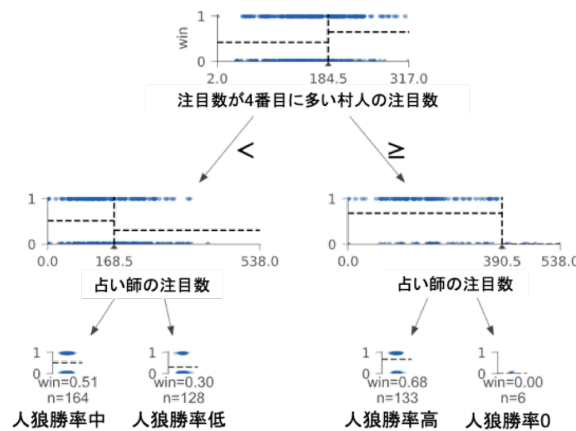


図 6: 発言数に基づく決定木 (ゲーム全体)

も目立たない」必要がある。発言の内容が有益かどうかを定量的に判断するのは難しい問題である。そのため文献 [3] では、目立っているかどうかに着目して分析を行った。目立っている程度を示す指標として、まず、発言が多いプレイヤーは目立っていると考え、あるプレイヤーの発言回数を発言数として用いた。もう一つの指標として、実際に他者の注目を浴びているかどうかを考慮し、あるプレイヤーが他のプレイヤーの発言で言及された回数を注目数として用いた。プレイヤーごとの発言数と注目数が勝敗にどう影響を与えるかを決定木を用いて調査した。16 人のキャラクターを属性、各キャラクターがゲーム全体を通して何回発言したかを属性値、ゲームの勝敗をクラスとして、発言数に関する学習データを作成した。注目数についても同様に作成した。また、収集した 431 件のデータにおいてゲーム最終日の中央値が 7 日目であったため、3 日目までをゲーム前半、4 日目以降をゲーム後半として分割し、それぞれの期間における発言数と注目数の学習データを作成した。この時、人狼や村人など複数存在する役職は、回数が多い順にソートした。また、キャラクターが 16 人に満たない村では、不足分をダミーの村人がいるものとして学習データを作成した。決定木の作成には python のライブラリである scikit-learn¹³を用い、決定

¹³<https://github.com/scikit-learn/scikit-learn>

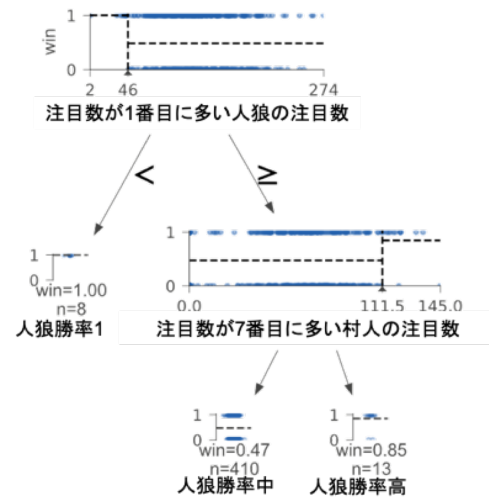


図 7: 注目数に基づく決定木 (ゲーム前半)

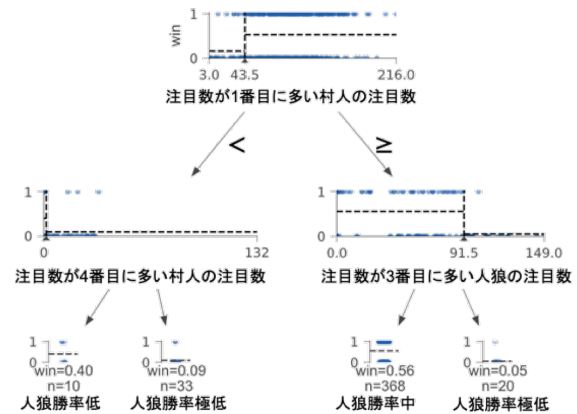


図 8: 注目数に基づく決定木 (ゲーム後半)

木の表示には dtreeviz¹⁴を用いた。

ゲームの全体、前半、後半の発言数と注目数による決定木を図 3 から図 8 にそれぞれ示す。図中の win は人狼陣営の勝率を表し、n は該当する村の数を表している。クラスである勝率を縦のパラメータとした。人狼陣営が勝った場合が 1、村人陣営が勝った場合が 0 に対応している。属性である各役職の発言数、および注目数を横のパラメータに対応させている。発言数、注目数について、ある閾値によって、分岐が生じる。右側の枝が閾値以上の場合に対応し、左側の枝が閾値未満の場合に対応する。最終段に表示されているのは、クラスの平均と、該当するクラスの個数であり、平均が高いほど、人狼の勝率が高いと言える。図 3 であれば、発言数が最も多い人狼の発言数が 54.5 以上であり、占い師の発言数が 40.5 未満である場合、人狼陣営の勝率が 71 ポイントと高く、該当する村が 9 件あったことを示している。leave-one-out 交差検証による分類精度を表 2 に示す。ゲーム前半の着目数についての分類精度が低く信頼性に乏しいが、他は 5 割以上の精度である

¹⁴<https://github.com/parrrt/dtreeviz>

表 2: 発話数と注目数に基づく分類精度

	全体	前半	後半
発話数	0.74	0.58	0.67
注目数	0.60	0.27	0.59

表 3: 囁き数に基づく分類精度

	全体	前半	後半
囁き数	0.62	0.58	0.57

ため、このデータに基づいて考察を行った。

発話数に基づく決定木を図3から図5に表示する。最初の分岐はどの決定木も人狼の発話数によってなされているため、人狼の発話数により勝敗が分岐することが確認できる。図3より、二段目の分岐にも人狼の発話数があり、人狼の発話数がさらに一定数以下であると人狼の勝率低下が見られる。また、このときの人狼の勝率は最低で9ポイントであり、発話数が最も多い人狼の発話数が閾値以上で占い師の発話数が閾値未満であれば、人狼の勝率は74ポイントとなり、65ポイント増加する。

ゲームの前半、後半における発話数に基づく決定木から、二段目の分岐には占い師の発話数があり、占い師の発話数が少ない方が人狼陣営の勝利に繋がる。つまり人狼が勝利するには初回から出来るだけ発話をし、占い師の発話数を制限する行動が必要である。

注目数に基づく決定木を図6から図8に表示する。注目数は特定の村人が人狼であると疑われているかどうか、占い師が本物であるかどうかの方に会話の主体があると思われる。図7より、ゲーム前半の注目数に基づく決定木をみると人狼の注目数は勝敗に関係しており、かつ一定数以下である場合、確定で勝利できる。しかし表2の精度を考慮すると、前半の人狼の注目数においてはあまり良い分類ではない。ゲームの後半に関しては、二段目に人狼の注目数による分岐がきているが、どちらに分岐しても勝率はそこまで高くない。つまり、ゲームの後半では人狼の注目数が低くても人狼の勝率陣営に有利であるとは言えない。以上から、発言数においては人狼の発話数が勝敗に関与しており、なおかつゲームの前半でもある程度の精度を持つので、この分析を基に特定のプレイヤーの発話数を操作する方法を考えていく。特に人狼が発話数を操作する際、他の人狼と作戦が被らぬよう、うまく連携する必要があり、またうまく連携が成されることで効率的に発話数を操作することが可能となる。人狼の連携について詳しくは6節で説明を行う。注目数においては思ったような成果は得られず、人狼への注目数が勝敗に与える影響があまり大きくない。特に注目数を減らす場合、発話数とは違い、処刑や襲撃などでゲームから除外しても制限することができない。そのため、注目数を上げないような発話内容や、役職の推定方法を今後考察していく。

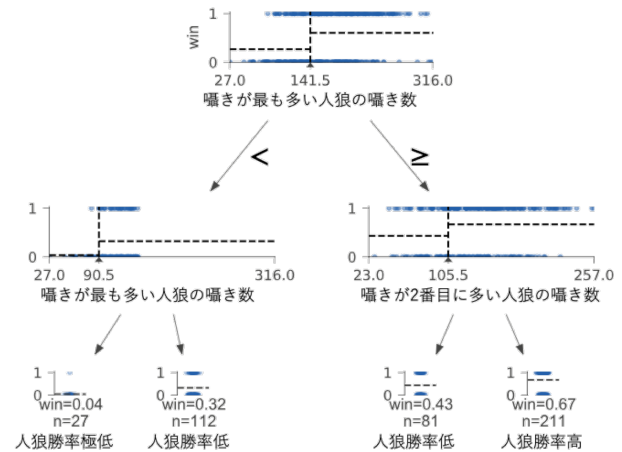


図 9: 囁き数に基づく決定木 (ゲーム全体)

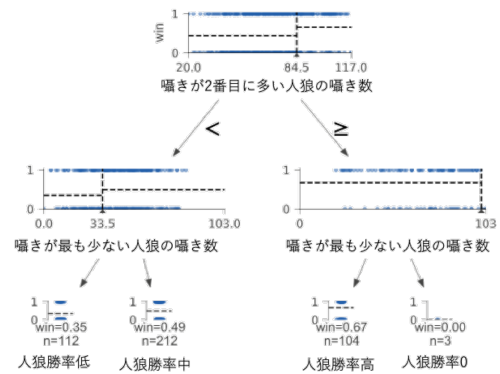


図 10: 囁き数に基づく決定木 (ゲーム前半)

6 人狼の連携度と勝率との関係の分析

6.1 囁き数

他の役職にはない人狼固有の能力として、4節で述べた「囁き」がある。囁きを利用することで秘密裏に人狼同士の連携が可能となり、勝敗に大きな影響を与えることができる。例えば、勝敗に影響を与える方法として、5節で述べたように、特定のプレイヤーの発話数を操作することが考えられる。特定のプレイヤーの発話数を操作する簡単な方法として、「問いかけ」がある。「問いかけ」とは、特定のプレイヤーに話しかけることで応答を募り、意図的に発話数を増加させることである。ただし一日にプレイヤーが発話できる数は有限であるため、一人で全てのプレイヤーの発話数を調整することは困難である。しかし自らの役職が人狼である場合には、囁きを介することで、人狼以外のプレイヤーに知られることなしに、発話数に余裕がある人狼に「問いかけ」をお願いし、人狼側が有利になるよう特定のプレイヤーの発話数を操作できる。各々の人狼が囁いた回数を囁き数として用い、人狼の囁き数が勝敗にどう影響を与えるかを決定木を用いて調査し

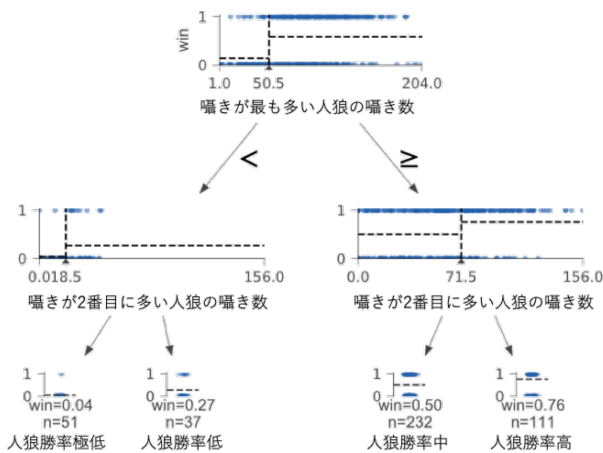


図 11: 囁き数に基づく決定木 (ゲーム後半)

た. 学習データの作成法は5と同様であり, 3人の人狼を属性, 各人狼がゲーム全体を通して何回囁いたかを属性値とした. 決定木の作成, 表示についても5と同様である. 囁き数による決定木を図9から図11にそれぞれ示し, leave-one-out 交差検証による分類精度を表3に示す. 属性である各人狼の囁き数を横のパラメータに対応させており, この決定木を基に連携度について考察する.

6.2 考察

囁き数が多いほど人狼同士の連携がとれていると単純にみなすことはできない. 例えば, 「問いかけ」や作戦の提案を行う際, 他の人狼が必ずしも自らの提案に乗ってくれるとは限らない. 自らの提案を受け入れてもらうには, 人狼ゲームの本質でもある対話を通して説得する必要がある. 対話をするということは, ある人狼の発言に対し他の人狼の応答がなされるので, 各人狼の囁きの数にはあまり差がない場合, 対話が成り立っていると考えられる. したがって, 人狼全体での囁きの数が多く, 各人狼の囁き数で差がないほど, 作戦や状況についてしっかり話し合い, 連携が取れていると仮定した. 図9より, ゲーム全体を通して, 囁き数が多い方が人狼の勝率は高い. 特に最初の分岐, 二段目の分岐ともに閾値以上である場合の人狼の勝率は67ポイントと高く, 最初の分岐, 二段目の分岐ともに閾値未満である場合に比べ, 勝率は63ポイントも上昇する. ゲームの前半では図10より, 最初の分岐は囁きが2番目に多い人狼であり, 閾値は84.5となっている. 二段目の分岐はどちらも囁きが最も少ない人狼によって成されているが, 最初の分岐で閾値未満の場合, 人狼の勝率は最大でも50ポイントである. 逆に最初の分岐が閾値以上の場合, 特異な三件を除いて人狼の勝率は高い. 最初の分岐において閾値以上の場合, 最も囁きが多い人狼もまた閾値以上の囁きを行なっているということであり, 少なくとも人狼二人は一定の対話を行なっていることが伺える. ゲームの後半では図11より, 最も囁きが多い人狼の囁き数で分岐が成されている. 二段目の分岐はどちらも二番目に囁きが多い人狼によって成されているが, 最初の分岐で閾値未満の場合, 人狼

の勝率は最大でも27ポイントであり, 高くはない. 逆に, 人狼の勝率が最も高くなるのは, 囁きが2番目に多い人狼の囁き数も多いときである. つまり, ゲームの後半においても対話を行い, 作戦についてよく話し合いをすることが勝率を上げる要因となる. まとめると, ゲームの前半でも後半でも最低人狼二人の対話が一定数が行われている場合, 勝率は高くなる. ただし, 今回の解析は人狼の連携度という観点では, 囁き数の差についての評価が成されておらず, 別の処理が必要となる. また, 今回収集した人狼が三人存在する村では, 囁きのデータ数が40MBと少ないため, データとしては不十分であり, 解析に使用する村の数を増やさなければならない. しかし人狼が三人ではない村も解析の対象に含める場合, 役職に変動があり, それにより採用する戦略も変わってくるため, 正規化を行う必要がある. つまり人狼の数, および役職に寄らない正規化の条件を決定することが当面の課題である.

7 まとめ

本稿では, 「人狼」エージェントに焦点をあて, 文献[3]で得られた, 一人の人狼の発言数が101を超えることで人狼陣営が勝利する可能性が最大65ポイント上がるという知見に基づき, 人狼同士の囁き数がゲームの勝敗とどのように関係するかを分析した. 囁きを解析することで最低でも人狼二人の囁き数が106以上の時, 人狼陣営が勝利する可能性が最大で63ポイント上がることを確認した. 今後は発言内容および囁きの内容も考慮し, 研究を進めていく. また, 阪本らが開発している人狼ゲームプラットフォームLiCOS[4]ではプレイヤーの本心と欺瞞を対話とは別に収集可能なため, 人狼の連携度を考慮する際に活用していきたい.

謝辞

この研究は平成30年度国立情報学研究所公募型共同研究の助成を受けています.

参考文献

- [1] 稲葉通将, 鳥海不二夫, 高橋健一: 人狼ゲームデータの統計的分析. ゲームプログラミングワークショップ2012論文集, pp. 144-147, (2012)
- [2] 稲葉通将, 大畠菜央実, 高橋健一, 鳥海不二夫: 雑談ばかりしていると殺される? 人狼ゲームにおける発言行為タグセットの提案とプレイヤーの行動・勝敗の分析. 情報処理学会論文誌, vol. 57, No. 11, pp. 2392-2402, (2016)
- [3] 永山翔滋, 阿部稔太郎, 大矢康介, 阪本浩太郎, 渋谷英潔, 森辰則, 神門典子: ステルス人狼エージェント作成に向けた発言数と注目数の分析. 言語処理学会第25回年次大会, P3-30, (2019)
- [4] 阪本浩太郎, 渋谷英潔, 森辰則: プレイヤによる他者の役職推定過程を記録する人狼ゲームプラットフォームLiCOSの開発と欺瞞コーパスの収集. インタラクティブ情報アクセスと可視化マイニング第21回研究会, (2019)