人狼ゲームにおける人狼陣営の勝率向上に寄与する行動分析

郡 涼太[1]　谷岡 広樹[1]

概要：近年，人工知能を活用したシステムが増加している．完全情報対戦型ゲームである将棋や囲碁でも，人工知能が活用されて名人を負かすまでになった．さらに，多人数不完全情報対戦型ゲームであるポーカーや麻雀においても人工知能は盛んに研究されている．一方，本研究で題材とする人狼ゲームは，不完全情報集団対戦型ゲームであり，不完全な情報を手がかりに集団としての勝利を目指す必要がある．本稿では，15 人制の人狼ゲームにおいて，人工知能がチーム戦略をとることで勝率が向上することを検証し，報告する．

キーワード：人狼ゲーム，チーム戦略, 人工知能

Behavioral analysis that contributes to improving the winning percentage of the werewolf team in the werewolf game

Abstract: In recent years, the number of systems that utilize artificial intelligence is increasing. Even in Shogi and Go, which are perfect information games, artificial intelligence has been used to beat masters. Furthermore, artificial intelligence is being actively researched in Poker and Mahjong, which are imperfect information games. On the other hand, the werewolf game focused in this study is an incomplete information collective game, and it is necessary to aim for winning as a team using incomplete information. In this paper, we report that artificial intelligence improves the winning percentage by taking a team strategy in a 15-player werewolf game.

Keywords: Werewolf Game, Team Strategy, Artificial Intelligence

1. はじめに

近年，人工知能を活用したシステムが増加している．完全情報対戦型ゲームである将棋や囲碁でも，人工知能が活用されて名人を負かすまでになった [1]．さらに，多人数不完全情報対戦型ゲームであるポーカー [2] や麻雀においても人工知能は盛んに研究されている．一方，本研究で題材とする人狼ゲームは，多人数不完全情報集団対戦型ゲームであり，不完全な情報を手がかりに集団としての勝利を目指す必要があるため，研究の余地がある．人狼知能大会 [3] は，人狼ゲームを人工知能だけで行う大会である．人狼知能大会では，人狼ゲームにおいて虚偽の発言や他エージェントの発言の真偽の判別精度が向上することで，人工知能の活用の可能性が広がると考えられる．この人狼知能大会では 5 人制と 15 人制があるが，5 人制では人狼が 1 人しかおらず，集団対戦ゲームとしての特徴が活かしきれないと判断し，本研究では 15 人制の人狼ゲームを対象とする．

1　徳島大学, Tokushima, Tokushima 770–8506, Japan

1.1　15 人制の人狼ゲーム

人狼ゲームは，プレイヤーが人狼陣営と村人陣営の 2 つの陣営に分かれて勝利を目指すゲームである．人狼陣営は，村人の人数を人狼と同数以下にすることが勝利条件であり，村人陣営は，人狼プレイヤーを全て処刑することが勝利条件である．15 人制の人狼ゲームでは，各プレイヤーを陣営（村人，人狼），種族（人間，人狼），役職（村人，占い師，霊媒師，狩人，人狼，裏切者）の属性を持つ．プレイヤーの属性は，人狼同士を除いて非公開情報である．

15 人のうち，村人 8 人，占い師 1 人，霊媒師 1 人，狩人 1 人の合計 11 人は村人陣営．人狼 3 人，裏切者 1 人は人狼陣営に属する．村人陣営の役職や能力については表 1，人狼陣営の役職と能力については表 2 にまとめる．ただし，裏切り者が生き残った際は村人陣営の勝利条件に加算される．

ⓒ 2021 Information Processing Society of Japan

表 1 村人陣営の役職の能力
Table 1 A List of Villager's Ability.

種族	役職	能力
人間	村人	なし
人間	占い師	夜のフェーズに 1 人の種族がわかる
人間	狩人	自分以外を 1 人襲撃から守ることができる
人間	霊媒師	前夜に追放したプレイヤーの種族がわかる

表 2 人狼陣営の役職の能力
Table 2 A List of Werewolf's Ability.

種族	役職	能力
人狼	人狼	夜のフェーズに全員で 1 人襲撃できる
人間	裏切者	なし

人狼ゲームには，昼と夜のフェーズが存在する．昼はプレイヤー同士で話し合いを行い，最後に投票を行う．投票で最多票を集めたプレイヤーが処刑される．夜では占い師が1人を占い，霊媒師が1人と交霊し，狩人が1人を守り，人狼が全員で1人を襲撃する．

1.2 人狼知能

人狼知能とは，人工知能の応用分野として発展途上といえる不完全情報ゲームの1つとして，人工知能研究者や情報系研究者に提案されたゲーム「汝は人狼なりや？」を人工知能のみで対戦を行う研究分野である [11].

本研究で使用する人狼知能のルールは 1.1 節のルールに加えて，会話についてのルールが追加される．プレイヤーの会話はターン制である．発言は定型文で 1 日に 10 回まで可能であり，発言をしないこと (SKIP，OVER) も可能である．全員が発言をするもしくは SKIP，OVER をすることを 1 ターンとし，これを繰り返す．発言する順番はターンごとにランダムに決定する．プレイヤーを P，役職を R，種族を S と表記して，定型文を表 3 に示す．

表 3 発言の定型文
Table 3 A List of Chatting Fixed Phrases.

発言	意味
COMINGOUT(P, R)	P は R であると宣言する
IDENTIFIED(P, S)	交霊の結果 P は S である
OVER()	発言終了
SKIP()	パス
VOTE(P)	P に投票する

2. 関連研究

多くの人狼知能の研究では，村人陣営の役職推定について注目したもの [4][9][10] が多い．この他，大川ら [5] は，ニューラルネットワークを用いて役職推定の精度向上を試みている．福田ら [4] は，会話情報を考慮しニューラルネットワークを用いて役職推定を行い，その有効性を確認している．伊藤ら [6] は，人間が用いる人狼ゲームの村人陣営の戦略を実装し，有効性を確認している．

3. 提案手法

関連研究では，村人陣営での自身への投票の誘導や役職推定など戦略の研究が盛んに行われているが，村人陣営のエージェント同士は味方か否かの判別ができないため，個人的な戦略しかとることができない．一方，人狼はゲーム開始時からお互いを把握しており，毎回の夜のフェーズに襲撃先決定前に人狼のみで会話することが可能なため，人狼は情報の非対称性を生かした戦略を用いることが可能になると考えられる．陣営ごとに共通して得られている情報の比較を表にしてまとめると表 4 のようになる．

表 4 情報の非対称性
Table 4 Information Asymmetry.

陣営	他エージェントの情報	味方エージェントのみで会話
村人	なし	不可
人狼	他の人狼情報あり	夜のフェーズで可能

そこで本研究では，チームとしての行動が可能な人狼に着目し，情報の非対称性を生かした戦略を適用する．人狼は，投票先に宣言したエージェントに対して同調して投票することで，村人陣営の票が割れた際に，人狼が投票するエージェントを追放しやすくする．ただし，仲間の人狼に票が集まる場合でも，過半数を超えて票が集まっている場合は，他の人狼の発覚を避けるために村人陣営の投票に同調するものとする．

また，人狼にとって直接的に発見されることが最も危険であるため，カミングアウトで占い師を宣言しているエージェントを人狼同士で意見を合せて，優先的に襲撃することで，占いによって役職が公表される危険を減らし追放されにくい状況をつくる．以上 2 点の戦略を人狼知能大会 2019 年優勝の takeda チームの人狼プログラム [7] に実装し，改良プログラムとする．

3.1 開発環境

本実験では，人狼知能大会のプラットフォーム aiwolf-ver0.6.2 [8] と aiwolf-docs-ver0.5.1 [8] を用いる．開発環境は eclipse，Java 環境の設定には JDK11 を使用した．

3.2 投票先決定の戦略

1 つ目の戦略は，投票時に人狼陣営で同調する戦略である．2019 年優勝プログラムの投票先決定アルゴリズムは，図 3.2 に示すように，投票先として最も多く宣言されたエージェントに対して投票を行う．

改良プログラムでは，図 3.2 に示すように，生存しているエージェントの数の半分未満の投票数の場合で，村人のエージェントが投票先として選ばれている場合は，人狼は同調して同じエージェントに投票を行うよう改良する．

Vol.2021-ICS-202 No.3
2021/3/29

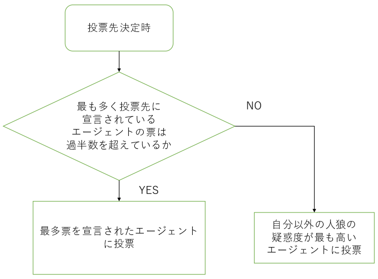

図 1　2019 年優勝プログラムの投票先決定プロセス
Fig. 1　Voting decision process of the 2019 champion.

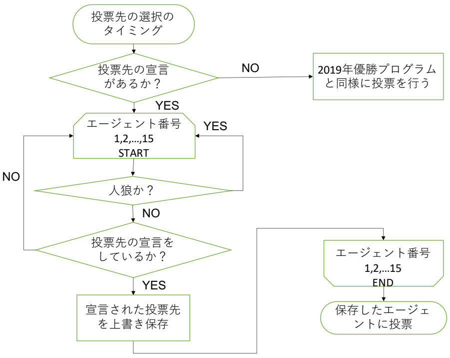

図 2　提案手法の投票先決定プロセス
Fig. 2　Voting decision process of the proposed method.

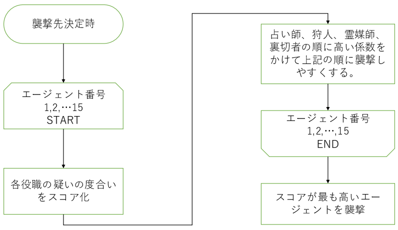

図 3　2019 年優勝プログラムの襲撃先決定のプロセス
Fig. 3　Target determining process of the 2019 champion.

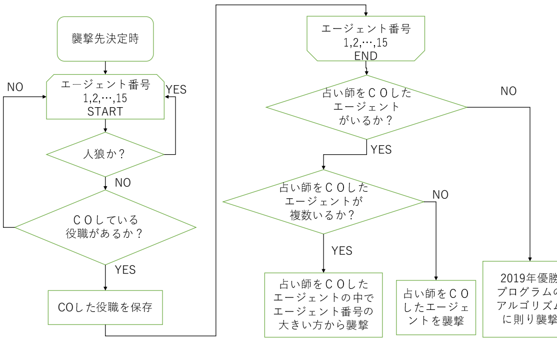

図 4　提案手法の襲撃先決定のプロセス
Fig. 4　Target determining process of the proposed method.

3.2.1　襲撃先決定の戦略

2 つ目の戦略は，襲撃時に占い師をカミングアウトしたエージェントを優先的に襲撃する戦略である．2019 年優勝プログラムの襲撃先決定アルゴリズムは，図 3.2.1 に示すように，占い師，狩人，霊媒師，裏切者の役職に対して，それぞれ疑わしさのスコアに，占い師，狩人，霊媒師，裏切者の順に優先度の係数を掛けたスコアを計算し，最も高いスコアのエージェントに対して襲撃を行う．

改良プログラムは，図 3.2.1 に示すように，人狼以外の全てのエージェントで占い師であるとカミングアウトしてるエージェントを襲撃先候補とする．占い師をカミングアウトしているエージェントが複数ある場合は，エージェント番号が大きい方を優先的に襲撃する．占い師をカミングアウトしているエージェントがいない場合は，優勝プログラムの襲撃先選択アルゴリズムに則る．

4.　実験

本研究では，人狼知能大会の 2019 年優勝プログラムと，人狼の行動のみ改良した改良プログラムを用いる．2019 年優勝プログラムと改良プログラムの対戦を行うことで，提案手法が勝率の向上に寄与したか否かを評価し，その結果を考察する．

4.1　実験概要

本実験では，2019 年優勝プログラムに，他のエージェントの投票先に同調する投票戦略と，占い師をカミングアウトしたエージェントを優先的に襲撃する襲撃戦略の 2 点を変更・追加した改良プログラムを作成し，村人陣営と人狼陣営で 100 回ずつ対戦を行った結果を集計し評価を行う．

4.2　実験結果

実験結果から，各プログラムの人狼陣営の勝利数を表 5 にまとめる．

表 5　対戦結果
Table 5　Match Results.

	2019 年優勝（村人）	改良（村人）
2019 年優勝（人狼）	36/100	29/100
改良（人狼）	59/100	58/100
投票戦略（人狼）	20/100	22/100
襲撃戦略（人狼）	55/100	58/100

5.　考察

人狼陣営の勝率の変動が誤差の範囲か否かを判断するため，カイ 2 乗検定を用いる．

5.1　定量的評価

本実験では，2019 年優勝プログラムと改良プログラムの勝ち数と負け数に差はない（帰無仮説）とし，2×2 分割表を対象にカイ 2 乗検定を行う．実測値を m，期待度数を e

ⓒ 2021 Information Processing Society of Japan

としたとき，カイ 2 乗値 χ^2 は式 (1) で求められる．

$$\chi^2 = \sum((m-e)^2/2) \quad (1)$$

自由度は 1．有意水準を 5% としたときの正規分布のカイ 2 乗値はカイ 2 乗分布表より 3.84 である．

まず，投票戦略と襲撃戦略を変更・追加した改良プログラムについて，勝利数をまとめると，表 6（() 内は期待値）となる．このとき，カイ 2 乗値は $\chi^2 = 10.60$ となる．

表 6　勝利数の比較
Table 6　Comparison of The Number of Wins.

	勝ち	負け
2019 年優勝	36(47.5)	64(52.5)
改良	59(47.5)	41(52.5)

$10.60 > 3.84$ であるため，帰無仮説は棄却される．なお，p 値は $p = 0.001$ (< 0.05)，効果量は $\phi = 0.23$ のため，有意差があると判断できる．

次に，投票戦略のみを変更・追加した改良プログラムについて，勝利数を纏めると表 7 (() 内は期待値)．このとき，

表 7　勝利数の比較（投票戦略）
Table 7　Comparison of The Number of Wins(Voting).

	勝ち	負け
2019 年優勝	36(28)	64(72)
投票戦略	20(28)	80(72)

カイ 2 乗値は $\chi^2 = 6.35$ となる．$6.35 > 3.84$ であるため，帰無仮説は棄却される．なお，p 値は $p = 0.012$ (< 0.05)，効果量は $\phi = 0.18$ のため，有意差があると判断できる．

さらに，襲撃戦略のみを変更・追加した改良プログラムについて，勝利数を纏めると表 8（() 内は期待値）となる．このとき，カイ 2 乗値は $\chi^2 = 9.71$ となる．$9.71 > 3.84$ で

表 8　勝利数の比較（襲撃戦略）
Table 8　Comparison of The Number of Wins(Voting).

	勝ち	負け
2019 年優勝	36(45.5)	64(54.5)
襲撃戦略	55(45.5)	45(54.5)

あるため，帰無仮説は棄却される．なお，p 値は $p = 0.002$ (< 0.05)，効果量は $\phi = 0.22$ のため，有意差があると判断できる．

5.2　定性的評価

投票先決定プロセスの改良後は，勝率が下がっている．エージェント番号の最も大きなエージェントが投票先に宣言したエージェントに投票する傾向が強く，この戦略が勝率向上にはつながらなかった．襲撃先決定プロセスの改良後は，有意に勝率は上がっており，カミングアウトされた役職が占い師であるエージェントを優先的に襲撃することで，襲撃の成功率が上昇した．

6.　結論

本稿では，情報の非対称性を活かしたチーム戦略を導入することで，人狼陣営の勝率を向上することを目指した．具体的には，2019 年人狼知能大会の優勝エージェントを改良し，村人陣営と人狼陣営を対戦させることで勝率を確かめた．実験結果から，人狼陣営が襲撃する対象を統一し，占い師を優先的に襲撃することで，有意に勝率が向上することがわかった．今後，他のエージェントにも有効かについて実験する必要がある．

謝辞

本研究は JSPS 科研費 JP18H03344 の助成を受けたものです．また，様々なご助言をいただいた徳島大学 A6 研究室のメンバーの皆さんに心より感謝申し上げます．

参考文献

[1] David S., Julian S., Karen S. et al.: *Mastering the game of Go without human knowledge*, Nature Vol.550, pp.354–359 (2017).
[2] Noam B. and Tuomas S.: *Superhuman AI for multi-player poker*, Science, Vol.365, Nov.6456, pp.885–890 (2019).
[3] 人狼知能プロジェクト，入手先 ⟨http://aiwolf.org⟩ (2021.02.12).
[4] 福田宗理, 穴田一: 15 人人狼ゲームにおける会話情報による役職推定, 2020 年度人工知能学会全国大会 (第 34 回) (2020).
[5] 大川貴聖, 吉仲亮, 篠原歩: 深層学習を用いて役職推定を行う人狼知能エージェントの開発, ゲームプログラムングワークショップ 2017 論文集, pp.50–55 (2017).
[6] 伊藤幹太,Reijer Grimbergen: 人間同士の人狼ゲームで用いられる戦術を反映させた人狼知能の研究, ゲームプログラムングワークショップ 2017 論文集, pp.183–188 (2017).
[7] 第 1 回人狼知能国際大会ソースコード (takeda チーム)，入手先 ⟨http://aiwolf.org/archives/2268⟩ (2021.02.12).
[8] 人狼知能プラットフォームの起動，入手先 ⟨http://aiwolf.org/server⟩ (2021.02.12).
[9] 池尚子, 三品晟瑠, 山根健: 人狼ゲームにおけるプレイヤの推測発言からの発言者および対象者の役職推定, 2020 年度人工知能学会全国大会（第 34 回）(2020).
[10] 早稲田凌亮, 中原航大, 渕田孝康: 深層学習による役職推定を用いた人狼知能エージェントの研究, 令和 2 年度電気・情報関係学会九州支部連合大会 (2020).
[11] 狩野芳伸, 大槻恭士, 園田亜斗夢, 中田洋平, 箕輪峻, 鳥海不二夫: 人狼知能で学ぶ AI プログラミング, マイナビ出版, pp.26–34 (2017).

ⓒ 2021 Information Processing Society of Japan