

5 人々狼における戦略進化のシミュレーション Simulation of Strategic Evolution in 5-player Werewolf

武田 惇史
Atsushi Takeda

鳥海 不二夫
Fujio Toriumi

東京大学
The University of Tokyo

In recent years, the game “Werewolf” has been drawing attention in the field of Artificial Intelligence. In this paper, we propose a method to simulate the evolution of strategy in order to investigate whether a strategy that always dominates other strategies is discovered in the 5-player werewolf regulation in future AI Wolf Competition. As a result of the simulation, we found that we can not reach a strategy that always has a dominant advantage, and eventually the strategy will continue to change periodically.

1. はじめに

近年, チェス, 将棋, 囲碁といった完全情報ゲームにおいて, 人工知能技術は人間の能力を上回るまでに成長している. それに比べると, 不完全情報ゲームにおける人工知能に関する研究はあまり行われてはいない. そこで, 人工知能研究者が次に取り組むべき題材として, 「人狼」が提案されており[篠田 14], 人狼における人工知能技術の促進のため, 統一的なプラットフォームの開発, 大会の開催などを行う「人狼知能プロジェクト」が発足している.

本研究の目的は, 人狼知能大会に用いられる「5 人々狼」を対象としたシミュレーション手法を提案し, 人狼知能においてナッシュ均衡となるような戦略が発見されるのかを調査することである.

2. 関連研究

人狼の元となったゲーム“Mafia”というゲームを対象にし, 役職として「人狼」と「村人」のみが存在する条件下で, 確率的統計より最適戦略を求める研究が行われている[Braverman 08]. ここでは, 話し合いをすることで, 村人勝率が上がることなどが示されている.

本研究が対象とする 5 人々狼において考えられる戦略については, 占い師の結果騙りが 5 人々狼において有効であるという, 一般的な人狼にない, 非直感的な戦略性について報告されている[片上 15].

人狼における進化的シミュレーションを行った研究[大澤 17]では, 人狼, 占い師, 村人の 3 人で行う 3 人々狼を定義し, 進化シミュレーションを行うことで, どのような戦略があるか, どのような現象がおこるかが調べられている. ここでも, 占い師の結果騙りが重要な戦略として挙げられていて, 進化の中で一時的に支配的な戦略になる, という結果が報告されている.

3. 5人々狼

3.1 ルール

人狼ゲームは, プレイヤーが「村人陣営」と「人狼陣営」に分かれて行うチーム戦である. 村人陣営は人狼を追放することで

表 1: 5 人々狼における役職の内訳

役職名	陣営	人数	能力
村人	村人	2	なし
占い師	村人	1	占い
人狼	人狼	1	襲撃
狂人	人狼	1	なし

勝利となり, 人狼陣営は人狼が追放されないことで勝利となる. ゲームは「日」を単位として進行し, 1 日はさらに「昼」と「夜」に分けられる.

本研究が対象とする 5 人々狼のレギュレーションにおいては, 以下のような順番でゲームは進行していく.

(1) 役職の割り当て

最初に, プレイヤーに対し役職がランダムに割り当てられる. 役職の内訳は表 1 に示す通りあらかじめ決まっている. また, 役職ごとに陣営が決まっている. プレイヤーは, 自分以外のプレイヤーの役職を知ることができない. したがって, 誰が敵で誰が味方かわからない状態でゲームが始まる.

(2) 0 日目: 占い師の占い

占い師は, プレイヤーを一人選んでその役職が人狼であるかどうかを知ることができる. この能力を「占い」と呼ぶ.

(3) 1 日目: 話し合い, 投票, 追放

ここでは, プレイヤーは話し合いを行う. 主に, 占い師 CO (役職を表明すること) と占い結果の報告が行われる.

その後, 各プレイヤーは他プレイヤーに対して投票を行い, 最も得票数の多かったものは「追放」される. 追放されたプレイヤーは今後ゲームに関わることはできない.

(4) 1 日目: 襲撃, 占い

ここでは, 人狼がプレイヤーを自由に選び, 「襲撃」を行う. 襲撃されたプレイヤーは今後ゲームに関わることはできない.

(5) 2 日目: 話し合い, 投票, 追放

この時点で残っているプレイヤーは 3 人である. 残ったプレイヤーに対して, 1 日目と同様に話し合い, 投票, 追放が行われる.

(6) 勝利陣営の決定

人狼が追放されたら村人陣営の勝ち, 最後まで人狼が生き残れば人狼陣営の勝ちである.

連絡先: 武田 惇史, 東京大学工学部システム創成学科,
113-8656 東京都文京区本郷 7-3-1 工学部 8 号館 526,
TEL: 03-5841-6991, E-mail: takeda@crimson.q.t.u-tokyo.ac.jp

表 3: 2017年データに対する回帰分析の結果と2018年における変化の対応

[illegible]

表 2: 特徵量

番号	概要	値
1	人狼占いCO	二値(0 or 1)
2	占い師2ターン目CO	二値(0 or 1)
3	2CO盤面における初日占い投票率	実数値[0, 1]
4	2CO盤面における二日目占い投票率	実数値[0, 1]
5	2CO盤面における占い結果騙り	二値(0 or 1)
6	2CO盤面における裏切り者黒だし	二値(0 or 1)
7	3CO盤面における占い結果騙り	二値(0 or 1)
8	3CO盤面における裏切り者對抗占い	二値(0 or 1)
9	3CO盤面における人狼対抗占い	二値(0 or 1)

3.2 考えられる戦略

人狼ゲームは、能力の弱い多数(村人陣営)と能力の強い少数(人狼陣営)の対戦であると解釈することができる。追放は、投票による多数決によって行われるため、村人陣営にとって有利であるといえる。このことが、人狼ゲームにおける「人狼陣営が村人陣営を騙る」という構図を作り出している。

5 人々狼においては、村人陣営の役職は村人または占い師であり、人狼陣営はこのどちらかを騙ることになる。どちらを騙るかは戦略において重要な要素である。また、占い師を騙る場合には嘘の占い結果を報告する必要がある。どのような嘘の占い結果を報告するかも、重要な要素になる。

4. 戦略の進化の分析

本章では、2017 年と 2018 年の人狼知能大会のログから、決勝に進出した強いエージェントのとりえる戦略がどのように遷移したのかを解析する。2018 年のエージェントは、2017 年において強かった戦法をとる傾向にあると予想される。

4.1 特徵量

解析に使う特徴量は表2に示す通りである。

特徴量の選択においては, 以下の条件を満たすものとした.

- 行動主体となる役職が存在すること。例えば、特徴量 1 は「人狼占い CO」であり、行動主体は人狼である。
- 行動主体の意思で値を変えられること。例えば、「村人が人狼に投票する確率」などは本分析では特徴量になり得ない。なぜなら、村人にとって誰が人狼であるかは非公開の情報であり、村人が自由にこの量を変化させることができないからである。

4.2 解析手法

試合ごとの特徴量の値を説明変数として、その試合の勝敗を2値(村人陣営が勝ちなら1, そうでなければ0)で予測するロジスティック回帰モデルを作り、学習した。回帰係数は、正負が逆であるという帰無仮説のもと、有意水準 1 パーセントで検定をおこなう。P 値が 0.01 を上回った特徴量は結果の考慮に入れない。

全データに対する特徴量の平均値の 2017 年から 2018 年における増減と、回帰係数の正負から、2017 年に対して強い戦略が 2018 年において用いられる傾向にあるかを明らかにした。

4.3 解析結果

結果を表3に示した。9個の特徴量のうち、8個でその回帰係数が有意であり、うち7個で回帰係数によって示されたエージェントにとって有利となる行動の増減と2017年から2018年における増減が一致した。このことから、2018年大会のエージェントは2017年大会に対して勝率が高くなるように作られていると考えられる。

5. シミュレーション

5.1 概要

第 4 章の結果に基づいて、ある年のエージェントは前年の戦略から学習するという仮定で、戦略の進化をシミュレートする。少ないパラメータで、行動原理に基づいた方法で妥当な結果を得ることを目指す。

シミュレーションでは、占い師、人狼、裏切り者のとる戦略を表す変数をそれぞれ S, W, P と書く。ありえるすべての状態に対して、その発生確率はこれらの変数で表される。村人は、 S, W, P の値を知ることはできないが、セオリーとしてそれぞれ S', W', P' に等しいという考えを持つものとする。このセオリーは、前年大会から学習したものと考える。 S', W', P' の値と各プレイヤーのとった行動から、ベイズ推定的に役職を推定することができる。プレイヤーの行動は、 S, W, P および S', W', P' を用いて計算される各プレイヤーの役職の確率分布によって決定される。起こりうるすべての状況に対しそれが起こる確率とその時の村人勝率の積の総和をとれば、村人勝率が分かる。すなわち、村人勝率は S, W, P, S', W', P' の関数として表すことができる。村人勝率を、 $p_{\omega}(S, W, P, S', W', P')$ と書くことにする。

各役職は、ほかの役職のとり戦略が変わらないという仮定のもと、自分の陣営の勝率を最大化するような値を探索する。

5.2 ゲームのモデル化

5人人狼モデルでは, 表 4 に示すような6つの変数を設定する.

以下に挙げるようないくつかの仮定においてゲームを単純化し、決定論的なシミュレーションを可能にする.

- 占い師 CO 者とその占い結果の組み合わせ以外の情報をプレイヤーが利用することはできない
- 2 日目に起こることは一切考えず、1 日目の時点で決着が付くものとして考える。
- 1CO 盤面になった場合、この時点で村人勝率は計算可能であり、それ以降のシミュレーションは行わない。

モデル化された5人人狼におけるゲーム進行を図 1 に示す。以下では、この図に従ってモデルの詳細を述べる。

(1) 占い師 CO

初日の最初のターンに占い師 CO が行われる。同時ターン制を仮定するので、人狼、裏切り者は互いの出方を見てから占い師を騙るかどうかを決定することができないものとする。

また、占い師は必ず占い師 CO するものとする。

人狼、裏切り者がそれぞれ占い師 CO をするかどうかで4通りの可能性がある。ゲームの進行上は、占い師 CO 者の人数ごとに3通りに分岐する。

(2) 1CO 盤面

1CO 盤面においては、村人勝率が 7/12 であることが解析的に求められる。

(3) 2CO: 占い師の結果出し

占い師の結果騙りは、村人を占ったとき占ったプレイヤー以外の人をランダムで選んで黒だしする、というパターンのみを考える。裏切り者、人狼に関しては、占い先としてありえるプレイヤー4通りと、占い結果(人狼か人狼でないか)の2通りで、 $4 \cdot 2 = 8$ 通りの可能性があり、その中から一様な確率でランダムに選んで行動する。

(4) 2CO: 勝利陣営の決定

1日目の投票が終わった時点で以下のように村人勝率を定めるとし、2日目以降のシミュレーションは行わない。

- 人狼が追放された場合、村人の勝ちとなる
- 裏切り者が追放された場合、確率 1/3 で村人の勝ちとなる
- 村人が追放された場合、村人の負けとなる。

(5) 3CO: 占い師の結果出し

プレイヤー5人のうち、占い師 CO した3人以外の2人は村人であることが全員の目線で確定する。占い師 CO した者たちの結果は、実質的に、村人に白出しするか対抗に黒出しするかのどちらかである。なぜならば、村人に黒出しすることは役職の割り当ての組み合わせが存在せず、論理破綻を起こすためである。また、それぞれの占い師 CO 者の目線において対抗の一人は裏切り者でもう一人が人狼であることはわかっているため、対抗の一人に白出しすることは、もう一人の対抗に黒出しすることと等価である。

人狼は、対抗の二人を区別する情報がないため、2人いる対抗を占う確率はそれぞれ $(1/2) \cdot w_2$ である。裏切り者についても同様のことが言える。また、真の占い師については、もし占い先が村人であった場合には確率 s_2 で対抗のどちらかに 50%の確率で黒出しするものとする。

(6) 3CO: 勝利陣営の決定

1日目の投票が終わった時点で以下のように村人勝率を定めるとし、2日目以降のシミュレーションは行わない。

- 人狼が追放された場合、村人の勝ちとなる。
- 裏切り者が追放された場合、確率 1/2 で村人の勝ちとなる。
- 村人が追放された場合、村人の負けとなる。

(7) 投票、追放先の決定

本モデルでは、村全体で合意の取れる人に最も投票が集まるものと考え、追放先を決定する。この仮定は、話し合いによって村人同士で投票先を合わせられること、人狼は2日目のことを考えて村人または占い師を騙らなければならない、怪しまれるような投票はできないことから、この仮定は妥当であると考えられる。

表 4: シミュレーション変数

変数名	意味
s_1	2CO 盤面における占い師の結果騙り率
w_1	3CO 盤面における占い師の結果騙り率
p_1	人狼の占い師CO率
s_2	3CO 盤面における人狼の対抗占い率
w_2	裏切り者の占い師CO率
p_2	3CO 盤面における裏切り者の対抗占い率

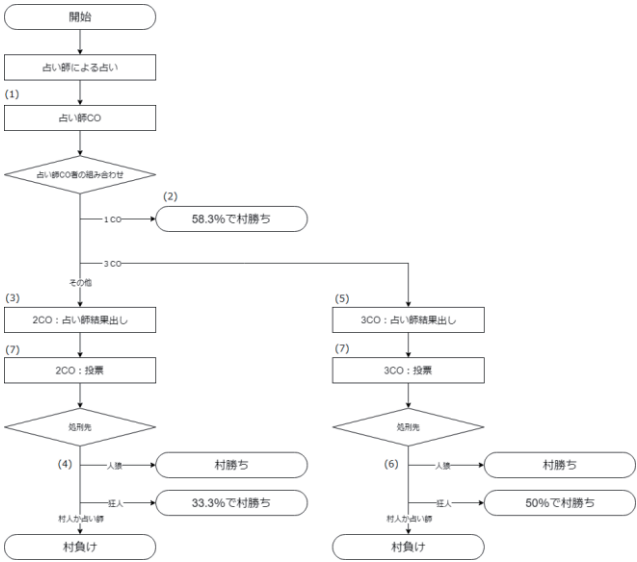


図 1: モデル化されたゲームのプロチャート

プレイヤーにとって、このモデルと S, W, P の値を既知とする場合、ベイズ推定によって各プレイヤー役職を推定することができる。実際には、村人たちは S, W, P の値を知らず、代わりに、第三者視点から、 S', W', P' を用いてベイズ推定により各プレイヤーの役職の確率分布を求める。村人たちはモデル自体を共有していることを仮定するので、勝率が最も高くなるような投票先を選ぶことが可能である。具体的には、3CO 盤面の場合、 $(人狼確率) + (1/2) \cdot (裏切り者確率)$ が最大のプレイヤーに投票し、2CO 盤面の場合、 $(人狼確率) + (1/3) \cdot (裏切り者確率)$ が最大のプレイヤーに投票する。

5.3 勝率計算

すべての変数の値を固定すれば、ゲームの考えるすべての展開をシミュレートすることによって、村人陣営の勝率の厳密な計算が可能である。すなわち、起こりうるすべての盤面の状態に対して、それが起こる確率とそのときの村人勝率の積の総和を計算すればよい。盤面の状態とは具体的に、3CO 盤面の場合、3人の占い師 CO 者の結果の組み合わせのことであり、2CO 盤面の場合、占い師を騙ったのが人狼と裏切り者どちらであるかと、占い結果の組み合わせのことである。

各盤面の状態に対する発生確率は S, W, P から計算可能である。そのときに追放されるプレイヤーは S', W', P' から計算可能であり、その時点で村人勝率は確定する。

以後、村人陣営の勝率を $p_v(S, W, P, S', W', P')$ と書くことにする。

5.4 シミュレーション条件

$s_1, w_1, p_1, s_2, w_2, p_2$ はそれぞれ 0.2 と 0.8 の二値を取るものとする。ある時点における流行りの戦略が (S', W', P') とあらわされ、それに対し、次のステップにおいて流行りの戦略が $(S_{new}, W_{new}, P_{new})$ になるとしたとき、

$$S_{new} = \operatorname{argmax}_S [p_v(S, W', P', S, W', P')]$$

$$W_{new} = \operatorname{argmin}_W [p_v(S', W, P', S', W, P')]$$

$$P_{new} = \operatorname{argmin}_P [p_v(S', W', P, S', W', P')]$$

が成り立つものとする。

5.5 結果と考察

シミュレーションの結果を図2に示す。ノードの書かれた数字は上段が左から s_1, w_1, p_1 の値であり、下が左から s_2, w_2, p_2 の値を表している。また、5.3 節における (S', W', P') から $(S_{new}, W_{new}, P_{new})$ に向かってエッジが付与されており、エッジを辿ることで戦略の遷移を追うことができる。

表3から、 s_1 が増加、 w_1 が減少、 p_1 が高い値を維持、 s_2, w_2, p_2 が低い値を維持している。ここで、2017 年、2018 年それぞれで採用されていた戦略を見ると、赤枠で囲われた箇所に対応していることが分かった。したがって、本シミュレーションは2017年から2018年の戦略の進化を内包しているといえる。

任意の状態から遷移をたどると、特定の1状態に収束せず、図2に青枠で示したサイクルに到達することが明らかとなった。従って、2017年及び2018年のデータとの対応が見られた赤枠の部分から遷移を辿っても、安定状態には到達しない。このことから、今後の人狼知能大会における戦略の推移において、ナッシュ均衡解は発見されない可能性が示唆された。

6. 結論

過去の人狼知能大会のデータログから、2017年、2018年大会においてどのような戦略が各陣営の勝利に貢献しているかを分析した。2018年のエージェントは2017年のエージェントに対して勝率が高くなるような戦略を取る傾向にあることを回帰分析により確認した。

また、5人狼のゲームを現実的な時間で網羅的、決定論的シミュレーションが行えるようにモデル化した。実際にシミュレーションを行うことで、遷移を繰り返していつでも特定の状態に収束せず、周期的に変化し続けるようになるという結果を得た。このことから、今後の人狼知能大会において、ナッシュ均衡解には辿り着かない可能性があることが明らかとなった。

本手法は、2日目を考えないなど大きな仮定を置いて、ゲームを単純化している。その点において、現実を正確に表せていないと考えられる。単純化を取り除くためには、行動原理の設定に工夫が必要である。しかし、どのように設定すべきかは非自明であり、より詳細なログ解析が必要である。この点に関しては今後の課題とする。

参考文献

[篠田 14] 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, 人工知能学会全国大会 JSAI2014, 2C4-OS-22a-3, 2014.

[Braverman 08] M.Braverman, O.Etesami, and E.Mossel. Mafia: A theoretical study of players and coalitions in a partial information environment. *The Annals of Applied Probability*, pp. 825–846, 2008.

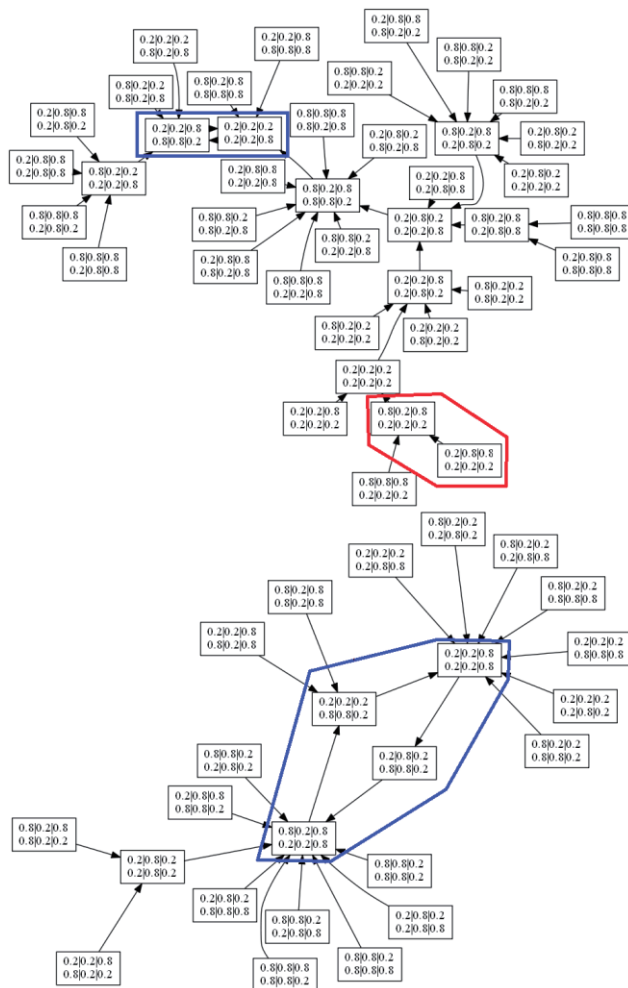


図 2: シミュレーション結果の遷移図

[片上 15] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 松原仁, 人狼知能プロジェクト, 人工知能学会論文誌, vol.30(1), pp.65-73, 2015.

[大澤 17] 大澤博隆, 汪博豪, 佐藤健: 進化シミュレーションを用いた3人狼の分析, 日本ソフトウェア科学会第 34 回大会講演論文集, 2017.

付録: 人狼用語

(1) CO

カミングアウトの略。自らの役職を表明すること。

(2) 1CO, 2CO, 3CO

占い師 CO する人数が1人, 2人, 3人であるようなゲームの状態をそれぞれ 1CO, 2CO, 3CO と呼ぶ。

(3) 対抗

占い師 CO したプレイヤーに対し、他の占い師 CO したプレイヤーを対抗と呼ぶ。

(4) 白出し, 黒出し

占いの結果報告が、「人狼でない」の場合白出し、「人狼である」の場合黒出しと呼ぶ。