

対話のない人狼ゲームの戦略

畢 曉恒^{1,a)} 田中哲朗^{2,b)}

概要: これまでの人狼ゲームの戦略に関する研究の多くは村人と人狼しかいない場合を扱ってきた。占い師のいる先行研究もあったが、占い師が吊られた時に人狼でないことが分かるという異なるルールが適用されていた。本研究では、本来のルールの中で占い師と狩人を導入した上で、双方に制約を加え、最適の戦略を計算することを試みた。双方の最適戦略は純粋戦略ではなく、混合戦略となるため、 ϵ -ナッシュ均衡を計算するのに CFR+ アルゴリズムを用いる。この結果、ゲームに関していくつかの考察を得た。

Strategies of the werewolf game without conversations

XIAOHENG BI^{1,a)} TETSURO TANAKA^{2,b)}

Abstract: Many researches about strategies in the werewolf game treat the limited game variation in which only villagers and werewolves exist. Although there was a research that treats the game with the seer, a rule applied in that research is different from the rules in reality, which says the identity of dead people will be revealed. This paper tries to calculate the optimal strategy of the werewolf game where the seer and the guard are both exist. The rules applied in this research are based on the rules in reality with some limitations. We used the CFR+ algorithm which calculates the ϵ -Nash equilibrium because the optimal strategy for villager side and werewolf side are mixed strategies but not pure strategies. The results suggest that there exist counterintuitive actions in the optimal strategy.

1. はじめに

人狼ゲームは情報の獲得、情報の真偽判断、自然言語処理など、様々な分野と関連しているため、人狼ゲームに関して研究することが多くの分野に貢献できると期待されている。一方で、人狼ゲームは会話を取り除いても、ゲームとして成立していると考えられる。この会話による説得を取り除いた人狼ゲームの戦略を検討することは、会話を取り入れた場合の戦略のベースラインとして意味があると考えられる。

会話を取り除いた人狼ゲームに関する先行研究はいくつ

か知られている。文献 [1] は人狼ゲームのバリエーションである「Mafia」に対し、会話を封じた状況での両陣営の戦略及び勝率に関して研究している。人狼ゲームでは人狼は毎日、村人を襲撃できるということに加えて、誰が人狼かを知っている、「吊り」の対象から人狼を外すように議論をコントロールできるという点が有利であるが、この先行研究ではこの有利さを生かせないように暗号化プロトコルを用いた同時投票を用いて、「吊りの対象を候補者の中から完全に一様に選択することが可能な手続きがある」という仮定を用いている。

この仮定のもとで、文献 [1] では占い師が存在しないルールと存在するルールの 2 種類に関して解析をおこなっている。その結果、占い師が存在しないルールでは、村人の数が R で人狼の数が $O(\sqrt{R})$ の時に勝率のバランスが取れること、占い師が存在するルールでは、村人の数が R で人狼の数が $O(R)$ の時に勝率のバランスが取れるという結果を得た。ただし、この「占い師が存在するルール」では占い師を吊った後に、吊られたのが人か、人狼のどちらかが全

¹ 東京大学総合文化研究科
Graduate School of Arts and Sciences, The University of Tokyo
bi@tanaka.ecc.u-tokyo.ac.jp

² 東京大学情報基盤センター
Information Technology Center, The University of Tokyo
ktanaka@tanaka.ecc.u-tokyo.ac.jp

a) bi@tanaka.ecc.u-tokyo.ac.jp

b) ktanaka@tanaka.ecc.u-tokyo.ac.jp

員に分かるという本来のゲームにはない仮定が用いられている。

文献 [2] では、占い師が存在しないルールで村人はランダムに投票して、人狼同士は村人からランダムに投票する条件での勝率を、いくつかの具体的な初期条件 (村人、人狼の数) のもとで計算している。ただし、文献 [2] が対象としたのは占い師が存在しない簡素化したゲームであり、また、「人狼同士が互いに投票をしなくても疑われない」ことに関する考察は不十分であると考えられる。

本論文の中で、まず、大規模な展開型ゲームの ϵ -ナッシュ均衡を計算する CFR+ アルゴリズムに関して紹介する。その後、CFR+によって人狼ゲームの戦略を計算するための手法を提案する。そして、CFR+によって計算した実験や実験の結果を示し、結果に関する考察を説明する。最後に論文のまとめと今後の予定について記す。

2. CFR+

文献 [3] に提案された CFR+ アルゴリズムは、CFR(counterfactual regret minimization)[4] のバリエーションの 1 種である。展開型ゲームにおける情報は $[N, H, A, Z, P, u_i \in N, \mathcal{I}_i \in N]$ という集合で表示することができる。その中、 N がプレイヤーの集合を示す。 H はゲーム木の状態を表示するヒストリーの集合であり、 A はゲームの流れをコントロールするアクションの集合である。ここに、 ha とは h の子状態であり、 h の状態からアクション a を実行したヒストリーと定義する。そして、 h は h' の先祖であることを $h \sqsubset h'$ で示す。 Z は全ての葉ノード (ターミナルとも呼ぶ) の集合であり、全ての $z \in Z$ は子孫がない。また、 $P(h) \in N \cup c$ は $h \in H \setminus Z$ における次にアクションを実行するプレイヤーを指す。ここに、 c はチャンスプレイヤーであり、 c に関する確率は事前に確定されている。 $u_i(z)$ とは、ターミナル z に到達した時に、プレイヤー i にとっての利得である。最後に、 \mathcal{I}_i はプレイヤー i の IS (information set) の集合である。IS とはゲーム状態の中、特定のプレイヤーが見える情報によって分けられたヒストリーの集合であり、ヒストリー h が含まれている IS は I^h として表示される。

プレイヤー i の戦略集合は σ_i として表示され、 $\sigma_i(a|I)$ によって、プレイヤー i が戦略 σ に従う時に、 I の IS の状態でアクション a を選択する確率を示すことができる。そして、計算するために、 I の IS の状態では必ずアクション a を選択し、他のヒストリーで σ に従うことを $\sigma: I \rightarrow a$ という表示方法に定義する。

そこで、 $\pi^\sigma(z) \equiv \prod_{ha \sqsubset z} \sigma_{P(h)}(a|I^h)$ をプレイヤーが戦略 σ に従う時の、ターミナル z の到達確率と定義し、この到達確率をさらにプレイヤー i に対して、自分がコントロールした $\pi_i^\sigma(z)$ と他のプレイヤー (チャンスプレイヤーを含む) がコントロールした $\pi_{-i}^\sigma(z)$ に分けることができる。

上記の定義により、プレイヤー i が戦略 σ に従い、 I の IS でアクション a を選択した時の期待利得は式 (1) のように表示することができる。なお、 $\pi_i^\sigma(h, z)$ はプレイヤーが戦略 σ に従う時の、ヒストリー h からターミナル z の到達確率のうち、自分がコントロールした寄与分を表している。

$$v_i^\sigma(I, a) \equiv \sum_{h \in I} \sum_{z \in Z: h \sqsubset z} u_i(z) \pi_{-i}^\sigma(z) \pi_i^{\sigma: I \rightarrow a}(h, z) \quad (1)$$

期待利得により、 $\sigma^1, \dots, \sigma^T$ という戦略のシーケンスに従う時の regret 値 (a を選択しなかった時の後悔度) を式 (2) で計算することができる。

$$R_i^T(I, a) \equiv \sum_{t=1}^T v_i^{\sigma^t}(I, a) - \sum_{t=1}^T \sum_{a' \in A} v_i^{\sigma^t}(I, a') \sigma_i^t(a'|I) \quad (2)$$

CFR は regret 値を利用し、式 (3) によって regret を減らすように戦略を反復的に更新するアルゴリズムである。

$$\bar{\sigma}_i^t(a|I) = \begin{cases} \frac{(R_i^{t-1}(I, a))^+}{\sum_{a' \in A} (R_i^{t-1}(I, a'))^+} & \text{if } \sum_{a' \in A} (R_i^{t-1}(I, a'))^+ > 0 \\ \frac{1}{|A|} & \text{otherwise} \end{cases} \quad (3)$$

ここで、 $(x)^+ \equiv \max(0, x)$ と定義している。CFR+ アルゴリズムはそれに加えて式 (4) により計算される Q 値を導入する。

$$\begin{cases} Q^0(I, a) = 0 \\ Q^t(I, a) = (Q^{t-1}(I, a) + R^t(I, a) - R^{t-1}(I, a))^+ \end{cases} \quad (4)$$

Q 値はマイナスにならないため、式 (5) のように、 Q 値を利用して戦略を更新することで、 σ をより高速に収束させることができる。

$$\bar{\sigma}_i^{t+1}(a|I) = \begin{cases} \frac{(Q_i^t(I, a))}{\sum_{a' \in A} Q_i^t(I, a')} & \text{if } \sum_{a' \in A} Q_i^t(I, a') > 0 \\ \frac{1}{|A|} & \text{otherwise} \end{cases} \quad (5)$$

このように、CFR や CFR+ アルゴリズムによって戦略を更新していくと、式 (6) による平均戦略が ϵ -ナッシュ均衡に収束する。

$$\bar{\sigma}_i(a|I) \equiv \frac{\sum_{t=1}^T (\sum_{h \in I} \pi_i^{\sigma^t}(h)) \sigma_i^t(a|I)}{\sum_{t=1}^T (\sum_{h \in I} \pi_i^{\sigma^t}(h))} \quad (6)$$

3. 提案手法

村人と人狼に加えて、占い師を導入する。文献 [1] のような占い師が吊られた場合は人狼でないという不自然な仮定を不要にするために、占い師を守る狩人も導入

する。他の役職(狂人, 霊能者等)は導入しない。

文献 [1] と同様に, 全員が同時に発言 (発言前に他の人の発言内容を参照しない) することが可能とする。

順番にしか発言できない状況でも以下のプロトコルにしたがうと同時発言を実現できる。

- プレイヤ i ($1 \leq i \leq n$) が自分の発言内容 m_i を決める。また乱数で文字列 r_i を生成する。
- プレイヤ 1 からプレイヤ n まで順に, 各プレイヤ i が十分な強度のある一方向ハッシュ関数 H を用いたハッシュ値 $H(m_i r_i)$ を発言する。
- プレイヤ 1 からプレイヤ n まで順に, $m_i r_i$ を発言する。

このプロトコルの成立には $H(mr) = H(m'r')$ となるペアを見つけるのが困難だという前提があるが, その前提を認めれば同時発言の実現法としては文献 [1] で提案されている方法よりも容易である。

同時発言を認めると, 先行研究にあるように, それぞれのプレイヤが 1-m までの数字を同時発言して, 総和の mod m をとることで, 1-m までのサイコロを振る (等確率で生成) ができる。これにより, 投票対象に関して合意が取れていれば, その中で等確率で投票ができる。

この条件の元で村人と人狼の戦略を考える。先行研究のように, 最適な戦略が導出できるわけではないので, 村人側, 人狼側のベースラインとなるように制約を加えた戦略について勝率を求める。

人狼の戦略のベースラインとして考えられるのが, 「ステルス人狼」という戦略である。一般の人狼ゲームでは, 人狼側が占い師等の役職者を騙ることをしないということになるが, 投票のコントロールもしないという状況の元では, 襲撃対象の選択だけが許された戦略となる。

人狼側が「ステルス」人狼という戦略を取るという前提の元では, 村人側は「占い師初日 CO(カミングアウト)」という戦略は有効となると考えられる。これに加えて, 以下の制約を加えた戦略を考える。

- 占い師が初日に CO(カミングアウト) する。

人狼が対抗 CO をしないという前提のもとで, 占い師の占いは常に信頼されるので, 占い師によって占われた対象が黒の場合はすぐに吊られる。また, 白の場合は白と占い師を除いた人 (以下では灰色と呼ぶ) の中からランダムに吊り対象を選択する。また, 占い師は吊り対象以外の灰色からランダムに占い対象を決定する。

狩人の行動に関しては,

- 占い師, 白, 灰色のうちのどれを守るか?

人狼の行動に関しては,

- 占い師, 白, 灰色のうちのどれを襲撃するか?

という選択肢がある。

これらは, 同じ状況 (同一の IS) で同じ行動を選択する (純粋戦略を採用する) のではなく, 確率的な行動を取る (混合戦略を採用する) 方が良いと予想される。

このようにモデル化したゲームは見かけは多人数だが, 理論的には 2 人ゼロ和ゲームであり, 混合戦略のナッシュ均衡が存在し, その利得が一意に決まるゲームとなる。このゲームの ϵ -ナッシュ均衡を求めるために, 前節の CFR+ を採用する。

4. 実験

文献 [3] に用いられる CFR+ のプログラムが公開されているが, ポーカーゲームに特化した内容であるため, 人狼ゲームに使うことができない。そのため, 本研究は python で汎用のプログラムを作成し, 同時着手にも対応する形で書き直した。実行の際は, 高速な Python 処理系である pypy を用いた。

本研究が用いるプログラムは文献 [3] の CFR+ プログラムと kuhn poker に関して同じ結果を得ることを確かめた。文献 [3] の CFR+ プログラムは論文本体の記述とは, 下記の点が異なっているので, それに従って実装した。これらは収束の高速化のための工夫であり, 本質的な違いはない。

- 戦略の更新は同じ t でもプレイヤごとに行う
- 平均戦略は t によって重みづけがされて, 新しいものほど重視される

本研究の実験では村人, 人狼, 占い師, 狩人の人数を変更して勝率, ϵ -ナッシュ均衡を与える混合戦略を求めた。こうすることで, それぞれの役目が勝率に与える影響や特定の状態での最優戦略を考察することができた。

作成した python プログラムでは, ゲーム状態 (ヒストリー) を下記のような文字列で表す

- 初期状態は「村人, 占い師, 狩人, 人狼」という順番に並ぶ 4 つの数字で表示する。ただし, 占い師, 狩人は 0 か 1 のみを許す。
- R は灰色の村人を指す
- D は占い師を指す (Mafia ゲームでの Detective)
- G は狩人を指す
- M は人狼を指す (Mafia ゲームからの借用)
- W は占い師の結果によって白確定になった村人を指す
- C は襲撃失敗により, 狩人視点で白確定と見なされている村人を指す。人狼視点でも C であることは分かる。
- S は占い師の結果によって白確定になった狩人を指す
- v は投票のステップを指す。v の次の文字は投票の結果吊られたプレイヤを表す。
- g は狩人の防御, 人狼の攻撃のステップを表す。
 - g の次の文字が狩人の防御対象を表す。狩人が既に死んでいる場合は「-」とする。
 - その次の文字が, 人狼の攻撃対象を表す。
 - 最後の文字が襲撃が成功したか (「+」), 失敗したか (「-」) を表す。
- d は占い師のステップを指す。

- dの次の文字が占われたプレイヤーを表す。 占い師がすでに死んでいるもしくは占った対象が襲撃された場合は「-」とする。

また、ヒストリーからゲームの状態及び狩人や人狼の視点で見える情報を同様の方法で表すことができる。

例として、下記のような表示は狩人の視点で、「村人四人、占い師一人、狩人一人、人狼四人の初期状態から、まず投票によって灰色のプレイヤーが吊られ、狩人と人狼両方も占い師を対象として指定したため、襲撃が失敗し、占い師は狩人を占い、投票で灰色のプレイヤーが吊られた」というISを意味する。

- G,4 1 1 4, vR, gDD-, dG, vR

同じヒストリーが人狼からの視点は下記のようなISになることがある。

- M,4 1 1 4, vM, gDD-, dR, vR

ステップの結果は違いがあるのは、狩人の視点から投票によって吊られたプレイヤーの身分を特定することができなく、逆に人狼の視点から狩人を村人の中から特定することができないためである。

実験の人数に関する設定は、下記の4つのグループに分けられる。

- (1) 占い師はいない, 狩人もいない
- (2) 占い師はいるが, 狩人はいない
- (3) 占い師はいないが, 狩人はいる
- (4) 占い師はいる, 狩人もいる

4つのグループの下で、CFR+を500回ループさせた村人側の勝率はそれぞれ表1, 2, 3, 4にまとめた。Texas Hold'emでは、exploitabilityの上限を計算するのが割合容易だが、今回のゲームではそれほど容易ではないため、exploitabilityは求めていない。そのため、現時点では500回ループ後の結果の精度が十分かどうかの保証はない*1。

総人数が10人以上に増えると、計算時間が長くなる。かつ、今回の目標は多人数の結果ではなく、異なる条件での結果の比較となるため、実験における総人数を10以上にならないと設定した。

5. 考察

実験の結果を考察してみると、いくつかのことが判明した。

5.1 占い師と狩人が勝率に与える影響

人狼が二人いる時の勝率比較は図1となる。単に占い師を入れると、初日は必ず襲撃される。最初の投票は占い師は対象外なので、人狼が吊られる確率がある程度あがるが、勝率への影響は大きくない。単に狩人を入れる場合も、勝率に与える影響は小さい。ただし、占い師と狩人を両方と

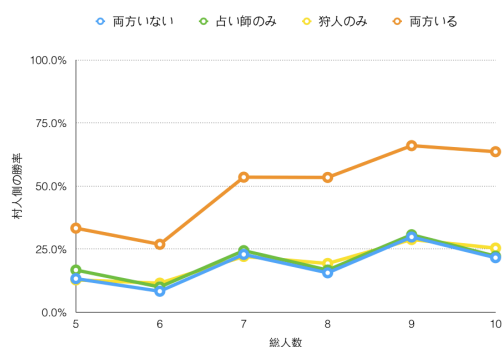


図1 人狼2人：異なる設定での勝率比較

Fig. 1 Mafia 2 : Comparison of winning percentage

も投入すると、影響が大きく、勝率が最大48%上がることもある。

5.2 狩人が占い師ではなく村人を守る方が有利な例

多くの場合、村人ではなく常に占い師を守るほうより利得が高いとわかった。しかし、白確定の村人がいなくても灰色の村人を守るほうが利得が高いケースもある。例えば、下記の場合は、人狼がまだ二人残っているなら、人狼の襲撃を失敗させるため、占い師を守るより、灰色の村人を守るほうが利得が高い。

- G,4 1 1 3, vR, gRR-, dG, vR, gDR+, d-, vC
- 平均戦略 R: 0.99, D: 0.01

5.3 狩人が人狼を守る方が有利な例

狩人の防御が成功したら、守る対象は狩人の視点から見ると白確定になるため、狩人の視点で残っている灰色の村人は全員が人狼だとわかれるケースがある。この場合、灰色を守る価値がなくなるのではないかと考えても不思議はない。しかし、すでに「残っている灰色は全員黒だ」とわかったとしても、わざと人狼を守るほうが利得が高いケースが見つかった。下記の場合は、人狼を守り、襲撃を成功させることで、人狼が投票で吊られる確率が上がるため、わざと灰色の人狼を守るほうが利得が高い。

- G,4 1 1 4, vR, gRD+, d-, vR, gRR-, d-, vC, gRR-, d-, vR, gRR-, d-, vR
- 平均戦略 C: 0, R: 1

5.4 人数による抽象化の妥当性

人狼ゲームの混合戦略を計算する時、総人数が10人以上の場合、人数が増えると、ヒストリーとISの規模は激しく大きくなるため、多人数のケースを計算する時、抽象化によってヒストリーとISの規模を小さくし、計算時間を短縮するという手法がありうる。抽象化として、ヒストリーを無視して現時点での自分の視点から見た人数だけを考えるのが自然な方法だと思われる。ただし、実験の結果

*1 一応、 $R_i^T \leq |Z|\Delta\sqrt{|A|T}$ という理論上の上限はある。

表 1 村人側勝率：占い師なし，狩人なし

Table 1 Winning percentage : no detective / no guard

人狼の数 \ 総人数	3	4	5	6	7	8	9	10
1	33%	25%	47%	38%	54%	45%	59%	51%
2	-	-	13%	8%	23%	16%	30%	22%
3	-	-	-	-	6%	3%	11%	7%
4	-	-	-	-	-	-	3%	1%

表 2 村人側勝率：占い師あり，狩人なし

Table 2 Winning percentage : with detective / no guard

人狼の数 \ 総人数	3	4	5	6	7	8	9	10
1	50%	33%	50%	40%	56%	46%	60%	51%
2	-	-	17%	10%	24%	17%	31%	22%
3	-	-	-	-	7%	4%	12%	7%
4	-	-	-	-	-	-	3%	1%

により，人数だけによる抽象化をおこなうと，最適な戦略を選択しなくなるケースがいくつか見つかった。

- M,5 1 1 1,vR,gRR-,dR,vC
- 平均戦略 W: 1, D: 0, R: 0
- M,5 1 1 1,vR,gDD-,dR,vR
- 平均戦略 W: 0.39, D: 0.61, R: 0.0

上記の例では，人狼の視点で，ヒストリーを無視すると，上と下は同じ IS になるが，戦略は大きな違いがあるため，人数による抽象化を用いると，最適な戦略ではなくなる。その原因に関して，ヒストリーを分析すると，上のほうは，「狩人が生きています」ということが保証されているが，下のほうは「狩人が死んだ」可能性があるため，抽象化がこの二つの場合を一緒にするため，不適切である。

- G,4 1 1 2,vR,gRR-,dR,vC
- 平均戦略 R: 0, W: 0.20, D: 0.80
- G,4 1 1 2,vR,gDD-,dR,vR
- 平均戦略 R: 0, W: 0, D: 1

狩人の視点でも，上記の例で，上のほうは推理によって最初の投票で死んだのは人狼とわかることができるが，下のほうはそれを判断することができないのは原因だと考えられる。前述のように，抽象化を用いてもより正確な結果を得るには，例のような区別を可能なように工夫する必要がある。

6. まとめ

今回の実験では，10人までの村に関して制限された行動のもとで，計算していくつかの考察を得た。今後はより多い村に関する計算していきたい。

本研究で用いたルールのもとで，狩人と人狼の実行できる行動を増やしたゲームを考えることができる。今回は狩人は自分が CO することはないという制約を加えたが，自分が吊られそうになった時や，かつて狩人が守ることに成功した村人が吊られそうになった時に，CO することは有効

な行動である。

また，人狼の行動に関しては，完全ステルスではなく，占い師騙りをする。狩人騙りをするようにすると選択肢が広がる。一方で，占い師騙りをする時，人狼側は

- 誰を占ったと騙り，その結果どうだったと騙るか？という行動で選択肢が広がり，村人側は
- 占い結果に応じて占い師も吊る (しかも確率的に)。という選択肢が広がる。狩人も
- 占い師を守る場合，どのような占い結果を残した占い師を守るか？という選択肢が広がるので，同じ人数でもゲーム木の規模は大幅に増える。そのため，また，最適な戦略が保証されないような抽象化も含めて有効な抽象化を考えると考えられる。

参考文献

- [1] M.Braverman, O.Etesami, E.Mossel. Mafia:A theoretical study of players and coalitions in a partial information environment. The Annals of Applied Probability, pp. 825846, 2008.
- [2] 西野順二, 自然な人狼の勝率. IPSJ SIG Technical Report, Vol.2015-GI-33, No.18.
- [3] M.Bowling, N.Burch, M.Johanson, O.Tammelin. Heads-up limit hold' em poker is solved. Science, vol.347, pp.145149,2015.
- [4] M.Zinkevich, M.Johanson, M.Bowling, C.Piccione. Regret Minimization in Games with Incomplete Information. NIPS-20, pp. 905912, 2008.

表 3 村人側勝率 : 占い師なし, 狩人あり

Table 3 Winning percentage :with guard / no detective

人狼の数 \ 総人数	3	4	5	6	7	8	9	10
1	33%	29%	46%	42%	53%	49%	58%	54%
2	-	-	13%	11%	22%	19%	29%	25%
3	-	-	-	-	5%	5%	11%	9%
4	-	-	-	-	-	-	2%	2%

表 4 村人側勝率 : 占い師あり, 狩人あり

Table 4 Winning percentage : with detective / with guard

人狼の数 \ 総人数	3	4	5	6	7	8	9	10
1	50%	33%	67%	68%	80%	81%	84%	83%
2	-	-	33%	27%	53%	52%	66%	64%
3	-	-	-	-	25%	20%	42%	40%
4	-	-	-	-	-	-	17%	15%