

# 人狼ゲームにおける同調行動の分析

郡田 和明<sup>1,a)</sup> 大久保 誠也<sup>2</sup> 西野 哲朗<sup>1</sup>

**概要:** 人狼ゲームとは、参加者全員で議論を行い、人狼だと思ふプレイヤーを多数決による投票で選ぶゲームである。多人数で議論を行うため、集団の意見や行動に沿うように個人の意見や行動を変容させる現象、同調行動が起こり得る。本研究の目的は、同調行動のメカニズムの解明である。そこで、人狼 BBS から得られたログを対象に、決定木分析を行った。分析の結果、投票先と占い先に関する議論で発生する「過半数の意見」とプレイヤーの意見が、一致しているか否かに着目することで、プレイヤーが再度意見をするとき、「過半数の意見」に同調するか反駁するかを 74% の的中率で予測できることがわかった。

## 1. はじめに

人狼ゲームとは、ゲーム参加者全員で議論を行い、人狼だと思ふプレイヤーを多数決による投票で選ぶコミュニケーションゲームである。多人数で議論を行うため、集団の意見や行動に沿うように個人の意見や行動を変容させる現象(同調行動)が起こり得る。近年、人狼ゲームをプレイする人工知能(人狼知能)を作る試み [1] に注目が集まっている。中村らは、人間同士の議論を再現できる人狼知能を作成し、人間が識別可能かを実験している [2]。このような人狼知能の課題として、篠田らは、他のプレイヤーへの同調・反駁機能の実装を挙げている [3]。先行研究として、稲葉らの研究 [4] がある。稲葉らは、人狼 BBS [5] のログを用いて、投票先と占い先に着目し、役職と同調・反駁の関係を分析した。結果として、人狼は人間と比べると、同調に対する反駁の割合が多いことがわかった。しかし、直前に発言したプレイヤー 1 人の意見に対する同調・反駁を扱っており、複数人の意見を扱っていない。

本研究の目的は、同調行動のメカニズムの解明である。そこで、決定木分析を用いた同調行動の予測を行うとともに、得られた決定木の分析を行う。これらの研究を通じて、人狼知能への同調行動機能の実装を目指す。本研究は、以下の手順により実施する。まず、人狼 BBS のログを収集し、投票先と占い先に関する発言のアノテーションを行う。次に、各ゲーム、日にち、プレイヤー毎に複数人の意見を考慮した「過半数の意見」を取得するとともに、「過半数の意見」

に対する同調行動を調べる。また、人狼ゲーム特有の同調行動の要因の検討を行う。最後に、要因を説明変数、同調行動を目的変数として決定木分析を行い、得られた決定木をロジスティック回帰分析等を用いて考察する。

## 2. 人狼ゲーム

人狼 BBS における人狼ゲームでは、ゲーム開始時に各プレイヤーに表 1 の役職のいずれかが与えられ、役職に応じて狼陣営と村陣営に分かれてプレイする。各陣営の勝利条件は、村陣営は人狼の数を 0 にすること、狼陣営は人間であるプレイヤーの人数を人狼の人数以下にすることである。

ゲームは昼と夜の 2 つのフェーズからなる。昼フェーズでは、誰が人狼であるかを探し出すための議論が行われる。このとき、役職を名乗り出たり、表 1 の能力を持つプレイヤーによる能力の使用結果が報告される。また、翌日の占い先や、今晚の投票先の公表などの、プレイヤー間の駆け引きが行われる。議論の後、すべてのプレイヤーは、人狼だと思ふプレイヤーを 1 人選び、投票を行う。全プレイヤーの投票によって、最多得票数を獲得したプレイヤーは処刑される。同数となった場合は、ランダムに処刑される。また、投票は無記名投票で行われ、プレイヤーは自身に投票することも可能である。処刑されたプレイヤーはゲームから除外され、ゲーム終了まで参加できない。夜フェーズでは、能力を持つプレイヤーには、能力に応じた情報が与えられる。人狼はプレイヤー 1 名を選択し、襲撃を行う。襲撃されたプレイヤーは、処刑されたプレイヤーと同様にゲームから除外される。

## 3. 分析データ

分析データの作成は、人狼 BBS のログの収集とアノテーションによるタグの付与、ならびに「過半数の意見」の取

<sup>1</sup> 電気通信大学大学院情報理工学専攻  
The University of Electro-Communications

<sup>2</sup> 静岡県立大学経営情報学部経営情報学科  
The University of Shizuoka

a) k1930045@uec.ac.jp

表 1 人狼ゲームにおける役職

役職	陣営	能力
村人	村陣営	なし
占い師	村陣営	夜フェーズでプレイヤーを1人指名し、そのプレイヤーが人間か人狼かを知ることができる
霊能者	村陣営	前日に処刑されたプレイヤーが、人間か人狼かを知ることができる
狩人	村陣営	夜フェーズでプレイヤーを1人指名し、そのプレイヤーを人狼の襲撃から守ることができる
人狼	狼陣営	夜フェーズでプレイヤーを1人指名し、襲撃することができる(人狼全体で1人襲撃)。また、他の人狼が誰であるか知ることができる
狂人	狼陣営	村人として扱われ、狼陣営勝利時に勝利となる

表 2 アノテーション結果の概要

アノテーションしたゲーム数	298 ゲーム
総発言	326136
アノテーションした発言数	21879

得からなる。まず、人狼 BBS のログの収集を行った。また、「過半数の意見」を考慮するにあたり、投票先と占い先に関する発言を取得する必要があるため、男子大学院生 1 人が人狼 BBS のログに対して、アノテーションによるタグの付与作業を行なった。アノテーション結果を表 2 に、アノテーション例を表 3 に示す。人狼 BBS では、重要な発言を行う際に、記号が用いられる。本研究では、アノテーションの都合上、投票先と占い先に関する発言は【】、▼、●等の記号を用いた発言に限定した。また、1 つの発言にタグは、複数個付与され得る。

「過半数の意見」の取得は、以下の手順で行った。

- (1) 同日発言内の、タグが付与されたプレイヤー X のある発言  $x$  とその直前のプレイヤー X の発言  $x-1$  を選択する
- (2) 発言  $x-1$  から発言  $x$  の間で、タグが付与された発言すべてを対象発言とする(発言  $x$  と  $x-1$  は含まない)
- (3) プレイヤー X を除く全プレイヤーの中で、対象発言内に含まれないプレイヤーが存在する場合、同日発言内で発言  $x-1$  までに付与された最新のタグを対象発言に加える(ただし、対象発言内に存在するプレイヤーはプレイヤー X を含めて 2 人以上とする)
- (4) 対象発言内で、プレイヤー毎の最新の発言を取得する
- (5) 得られたタグの集合から、最大要素数に該当するタグを「過半数の意見」とする

表 3 を用いて例を説明する。プレイヤー X をジムゾン、プレイヤー X のある発言  $x$  を log 番号 26、直前の発言  $x-1$  を log 番号 4 とする。このとき、「過半数の意見」の取得手順に従うと、対象発言は、log 番号 2,3,10,15 となり、タグの集合として { ジムゾン, ジムゾン, ペータ, ジムゾン } が得られる。最大要素数に該当するタグである「ジムゾン」が「過半数の意見」となる。最終的には、log 番号 26 のタグ「パメラ」と「過半数の意見」であるタグ「ジムゾン」を比較し、同調しているか反駁しているかを判定することになる。

同調行動は、同調や反駁することを意味し、同調・反駁の基準は、先行研究を参考にし、下記のように定義する。

**同調と反駁：** 発言 X が発言 Y (X と Y は異なる) に対して同調した発言であるとは、双方の付与タグが完全一致もしくは部分一致した場合のことをいう。反駁した発言であるとは、双方の付与タグが同調の条件を満たしていない場合のことをいう。

以上より得られた分析データの概要を、表 4 に示す。

#### 4. 同調行動の要因

同調行動の要因として、本研究では大まかに 4 つのカテゴリを検討した。基礎データ、議論の詳細なトピック、議論の状況、意見の分布の 4 つである。

1 つ目の基礎データは、先行研究や関連研究に基づいた要因である。先行研究では、投票先や占い先といった議論の話題と、人狼や人間といったプレイヤーの役職に着目している。人狼ゲームとは異なるが、心理学の分野において個人の専門性と同調行動の関係について調査した研究 [6] がある。そこで、本研究では、プレイヤーの熟練度も考慮する。プレイヤーの熟練度は、熟練者と初心者の 2 種類を用意し、人狼 BBS が初めてのプレイヤー、もしくは初めての役職を持ったプレイヤーを初心者とし、それ以外を熟練者としている。よって、議論の話題、プレイヤーの役職、プレイヤーの熟練度の計 3 種類の変数を基礎データとして扱い、ベースラインとする。

2 つ目の議論の詳細なトピックは、基礎データで扱う議論の話題と比べて、より詳細な議論の内容や雰囲気を表す。例えば、議論の長さを表す発言量や議論の活発さを表す他プレイヤーの発言の引用回数等を考慮する。また、LDA [7] を用いて、議論に含まれるトピックを求め、各トピックの含有率を要因に加える。LDA はトピック数と文章群を入力すると、文章群からトピックを抽出するとともに、各文章が含むトピックの含有率を計算する。LDA を用いることで、議論の話題だけでなく、霊ロラ等の戦術に関する詳細な話題を考慮できると考えられる。本研究では、得られるトピックの解釈性を上げるために、品詞の制限とトピック数 (2, 3, ..., 6) とストップワードを設定する。品詞の制限では、名詞の一般形や副詞等の計 15 種類の品詞に限定し、品詞の組み合わせとして 24 種類検討した。ストップワードには、出現頻度 (出現文書数/全文書数) が一定以上の割合 (0.1, 0.2, ..., 0.5) の単語を加える。品詞の組み合わせ毎に、適切なトピック数と割合の設定を行う。具体的には、学習データに対して LDA を用いてトピック含有率を求め、トピック含有率を説明変数、同調行動を目的変数として決定木分析を行い、10 分割層化交差検証で的中率を求める。分析は、python の gensim, StartifiedKFold, GridSearchCV, wordcloud を用いた。例として、品詞の組み合わせが名詞と人狼 BBS の専門単語、適したトピック数が 4、割合が 0.4 であった LDA から得られたトピックを図 1 に示す。図 1



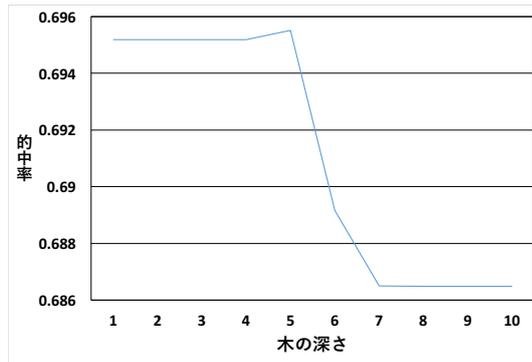


図 2 意見の分布を用いた決定木の木の深さと的中率の関係

表 8 compare\_opinion と同調行動のクロス分割表  
 $\chi^2 = 192.46, df = 1, p < .01$

compare_opinion	反駁	同調	合計
不一致	232	49	281
一致	141	326	467
合計	373	375	748

表 9 compare\_opinion と同調行動の残差分析  
 $*p < .05, **p < .01$

compare_opinion	反駁	同調
不一致	13.873**	-13.873**
一致	-13.873**	13.873**

表 10 ロジスティック回帰分析の混同行列

		正解	
		同調	反駁
予測	同調	299	119
	反駁	76	254

の分岐であるプレイヤー X の発言  $x-1$  と「過半数の意見」が一致しているか否か (compare\_opinion) によって予測結果が決まっていた。ゆえに、compare\_opinion と同調行動の関係を調べる。具体的には、カイ二乗検定と残差分析を行う。分析結果を表 8 と、表 9 に示す。

compare\_opinion が他のモデルでも有効に働くかを調べるために、表 5 のカテゴリ 1+2+3+4 を用いて、ロジスティック回帰分析を行った。分析は R の glm, step を用いた。多重共線性を考慮し、相関が 0.8 以上の変数は片方の変数を削除し、変数増減法を用いてモデルを構築した。ロジスティック回帰分析では、合計 26 個の変数が選択され、その中に compare\_opinion が含まれていることが確認できた。ロジスティック回帰分析の予測結果に対する混同行列を表 10 に示す。

## 5.2 完全一致と部分一致を区別した同調行動

5.1 節では、同調の条件として、タグの完全一致と部分一致を区別していなかった。本節では、完全一致と部分一致を区別した場合について、同様の決定木を用いた分析を行う。分析データを表 11 に示す。決定木を用いた予測結果を表 12 に示す。

表 11 学習データとテストデータ

データ	同調 完全一致	同調 部分一致	反駁件数
学習データ	842	857	859
テストデータ	224	209	207

表 12 予測精度の一覧

カテゴリ	木の深さ	学習データの 最終精度	テストデータの 的中率
(1) 基礎データ	1	0.410	0.406
(2) 議論の詳細な トピック	6	0.431	0.378
(3) 議論の状況	3	0.419	0.420
(4) 意見の分布	2	<b>0.727</b>	<b>0.742</b>
1+2	6	0.435	0.409
1+3	2	0.449	0.448
1+4	2	0.727	0.742
2+3	10	0.461	0.421
2+4	2	0.727	0.742
3+4	2	0.727	0.742
1+2+3	6	0.466	0.434
1+2+4	2	0.727	0.742
1+3+4	2	0.727	0.742
2+3+4	2	0.727	0.742
1+2+3+4	2	0.727	0.742

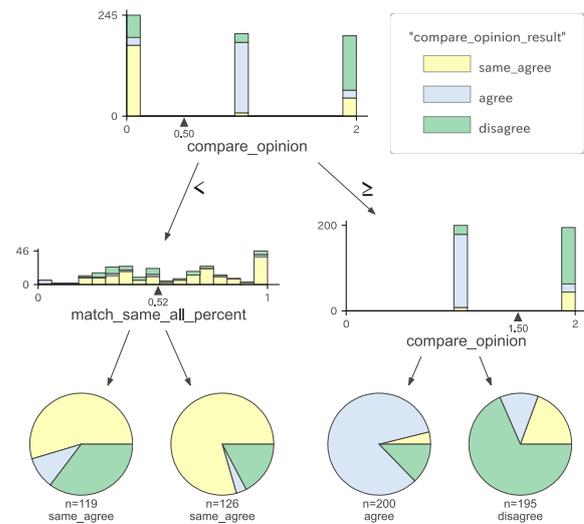


図 3 意見の分布の決定木

得られた決定木を図 3 に示す。compare\_opinion の 0 は完全一致、1 は部分一致、2 は不一致を表す。compare\_opinion\_result は同調行動を表し、same\_agree は完全一致、agree は部分一致、disagree は反駁を表す。また、match\_same\_all\_percent は全プレイヤーの人数に対するプレイヤーの  $x-1$  の意見と完全一致している人数の割合を表す。次に、パラメータチューニングの結果を図 4 に示す。

## 6. 考察

### 6.1 完全一致と部分一致を区別しない同調行動

表 5 をみると、基礎データを用いた場合は的中率が 0.574

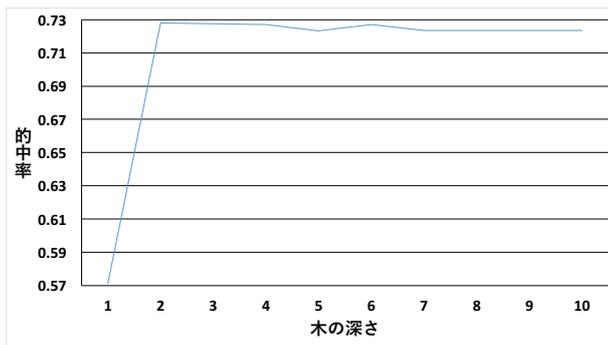


図 4 意見の分布を用いた決定木の木の深さと的中率の関係

であり、意見の分布を用いた場合は、的中率が0.745まで上昇している。得られた決定木の妥当性を考えると、図2より、木の深さを5より深くした場合、的中率が減少している。これは、決定木が学習データに対して過剰に適合したからだと考えられる。しかし、的中率は0.68から0.70の範囲で推移しているため、パラメータチューニングによって著しく的中率が変化することはないと考えられる。また、表5より、意見の分布を含んだカテゴリは木の深さが1であることから、ルートノードの分岐である compare\_opinion が同調行動の予測において重要な変数であると考えられる。よって、ルートノードの分岐 compare\_opinion について考える。表8では、compare\_opinion と同調行動の関連性をカイ二乗検定を用いて検証している。表8より、1%水準で compare\_opinion と同調行動に関連があることがわかった。そして、表9の残差分析より、プレイヤーの  $x-1$  の意見とプレイヤー以外の「過半数の意見」が完全一致もしくは部分一致のとき、反駁より同調が多いこと、プレイヤーの  $x-1$  の意見とプレイヤー以外の「過半数の意見」が不一致のとき、同調より反駁が多いことが明らかとなった。この理由を明らかにするために、意見の分布を用いた決定木を分析したが、有効な示唆が得られなかった。理由を考察する上では、意見の分布を含めたその他のカテゴリでも同程度の的中率が得られているため、今後、決定木の深さを深くし、分析することで考察できると考えられる。

一方で、compare\_opinion は、決定木を用いたため選ばれた変数である可能性が考えられる。大塚らの研究 [9] では、ロジスティック回帰分析と決定木分析の両モデルを用いて予測を行い、使用された共通の説明変数は目的変数の予測において頑健な変数であると評価している。そこで、ロジスティック回帰分析を行い、compare\_opinion が目的変数を説明する上で、モデルに関わらず、頑健な変数であることを確認する。ロジスティック回帰分析の結果、compare\_opinion は説明変数として選択されたため、頑健な変数であると言える。また、ロジスティック回帰モデルの的中率は0.742であり、決定木と比較しても予測精度に大差ない。しかし、ロジスティック回帰分析で採用された変数が決定木分析では採用されていない現象が確認できた。原因としては、決定木

表 13 表 11 の分析データの compare\_opinion と同調行動のクロス分割表

compare_opinion	同調行動		
	同調 完全一致	同調 部分一致	反駁
同調 完全一致	0.676	0.069	0.255
同調 部分一致	0.040	0.837	0.123
反駁	0.202	0.115	0.683

分析においてパラメータの調整を行ったことが挙げられる。

## 6.2 完全一致と部分一致を区別した同調行動

次に、同調の完全一致と部分一致を区別した場合の結果について考察する。表12をみると、基礎データを用いた場合は的中率が0.406であり、意見の分布を用いた場合、的中率は0.742まで大幅に上昇している。図4より、パラメータチューニングの結果、木の深さを2より深くすると的中率が上昇し、以降は一定の的中率を保っている。これは、木の深さを深くすることだけでは的中率が変化しないことを示唆している。次に、表12より意見の分布を用いたカテゴリは木の深さが2であることが、また、図3の可視化から、compare\_opinion の分岐によって学習データに対して最適化されていることがわかった。ゆえに、compare\_opinion の分岐の妥当性を検証する必要がある。カイ二乗検定を行った結果、ルートノードと右の子ノードにおける compare\_opinion と同調行動は1%水準で関連があることがわかった。

最後に、compare\_opinion をふまえた同調行動の推移に着目した考察を行う。表13は、行毎の割合を示したクロス分割表である。表13から、「過半数の意見」と部分一致したプレイヤーは、ほとんどが同調(部分一致)を行っていることがわかる。そして、「過半数の意見」と完全一致したプレイヤーは約3割が反駁し、「過半数の意見」と不一致のプレイヤーは約3割が同調(完全一致)しているという逆の現象が確認できる。まず、「過半数の意見」と完全一致したプレイヤーは約3割が反駁している現象については、match\_same\_all\_percent の割合が関係していると考えられる。match\_same\_all\_percent は、全プレイヤーの人数に対するプレイヤーの  $x-1$  の意見と完全一致している人数の割合である。図3より、左の子ノードに着目すると、match\_same\_all\_percent が0.52以上の場合は同調(完全一致)の割合が75%を占めており、0.52未満の場合は60%近くまで割合が減少していることが分かる。そして、match\_same\_all\_percent の割合が0.52未満の場合、反駁を行うプレイヤーが4割近く確認できるため、これが約3割の反駁を表していると考えられる。match\_same\_all\_percent が0.52未満という状況は、「過半数の意見」と自分の意見は完全一致しているが、「過半数の意見」と完全一致しているプレイヤーは自分以外は少ない状況を表す。つまり、「過

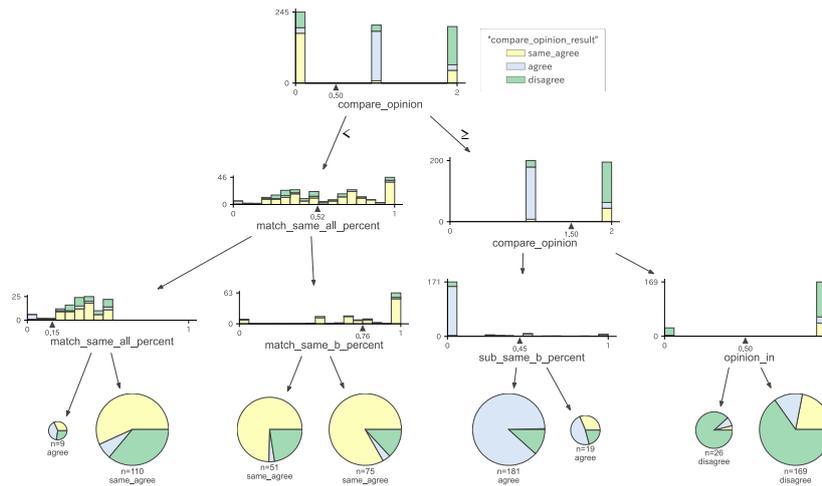


図 5 木の深さが 3 の意見の分布を用いた決定木

半数の意見」は様々な意見から構成されており、各プレイヤーの意見は統一されていない。ゆえに、議論は混沌としている可能性が高く、「過半数の意見」に反駁するような現象が起こり得ると考えられる。次に、「過半数の意見」と不一致のプレイヤーは約 3 割が同調 (完全一致) しているという逆の現象が確認できる。この理由としては、決定木の深さを深くし、得られたノードから考察ができると考えられる。これらの結果をふまえて、木の深さの範囲を (3, 4, ..., 10) に設定し、意見の分布を用いて再学習を行った。得られた決定木を図 5 に示す。図 5 から、「過半数の意見」と不一致のプレイヤーが、「過半数の意見」に自身が含まれていない場合 (opinion\_in の 1), 同調 (完全一致) の割合が 3 割程度を占めていることがわかった。これは、「過半数の意見」への同調 (完全一致) を検討する際に、「過半数の意見」に自身が含まれている場合は、最後まで「過半数の意見」に抵抗する必要があるが、自身が含まれていないため、「過半数の意見」に同調しようと考えたのではないかと考えられる。

## 7. おわりに

本研究では、同調行動のメカニズムの解明を目的とし、人狼 BBS のログを対象に、決定木分析を行った。本研究では、同調行動は、他者の意見に同調することや反駁することを意味し、他者の意見を複数人の議論における「過半数の意見」として取り扱った。

決定木分析の結果、投票先や占い先を決める議論の場で発生する「過半数の意見」とプレイヤーの意見が、“完全一致しているか部分一致しているか不一致なのか”に着目することで、プレイヤーが再度意見をするとき、「過半数の意見」に“同調するか反駁するか、同調する場合は完全一致するのか部分一致するのか”を約 74% の確率で予測することが可能であることがわかった。

また、カイ二乗検定と残差分析の結果から、 $x-1$  の意見とプレイヤー以外の「過半数の意見」が完全一致もしくは部

分一致しているとき、反駁より同調が多いこと、プレイヤーの  $x-1$  の意見とプレイヤー以外の「過半数の意見」が不一致のとき、同調より反駁が多いということがわかった。

今後の課題としては、研究成果の人狼知能への実装や、深層学習や複雑なモデル等を用いた予測精度の向上、投票先や占い先に関する発言に対するアノテーションの自動化などが挙げられる。

**謝辞** 本研究を進めるにあたり、人狼ゲームのログデータを使用する許可をいただいた人狼 BBS の管理者 ninjin 様に深く感謝し、御礼申し上げます。

## 参考文献

- [1] 鳥海不二夫, 梶原健吾, 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐, 西野順二, 人工知能は人狼の夢をみるか—人狼知能プロジェクト, デジタルゲーム学会, 2014.
- [2] 中村 洋文, 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 狩野芳伸, 人間同士の議論の疑似体験を実現する人狼知能対戦システムの開発, 第 33 回ファジィシステムシンポジウム 講演論文集 (FSS2017 山形大学).
- [3] 篠田 孝祐, 片上 大輔, 稲葉 通将, 鳥海 不二夫, 大澤 博隆, 松原 仁, 人狼知能プロジェクト 2015, 人工知能学会研究資料 SIG-SLUD-B502-20.
- [4] 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐, 鳥海不二夫, 議論の構造に着目した人狼ゲームの分析, The 19th Game Programming workshop 2014.
- [5] 人狼 BBS, <http://ninjinix.x0.com/wolf0/>.
- [6] 吉武 久美子 同調・非同調者の自己意見及び他者意見の想起に関する実験的研究, The Japanese Journal of Experimental Social Psychology. 1988, Vol. 27, No. 2, 163-169.
- [7] 奥村 学, 佐藤 一誠, トピックモデルによる統計的潜在意味解析, コロナ社, 2015.
- [8] 稲葉 通将, 狩野 芳伸, 大澤 博隆, 大槻 恭士, 片上 大輔, 鳥海 不二夫, 人狼 BBS に対する役職表明・能力行使報告情報のアノテーション, The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018.
- [9] 大塚 祐輔, 平間 一樹, 横田 賀英子, 渡邊 和美, 和智 妙子, 単発の殺人における犯人の犯罪経歴の予測手法-ロジスティック回帰分析と決定木の比較-, 法科学技術, 22(1), 25-34(2017).