

雑談ばかりしていると殺される？ 人狼ゲームにおける発話行為タグセットの提案と プレイヤーの行動・勝敗の分析

稲葉 通将^{1,a)} 大島 菜央実^{1,†1} 高橋 健一¹ 鳥海 不二夫²

受付日 2016年2月19日, 採録日 2016年9月6日

概要: 本研究では, 人狼ゲームにおけるプレイヤーの発話内容を表現するタグを設計し, それらのタグを人狼 BBS におけるプレイヤーの発話に付与し分析を行った. 分析では, 襲撃対象, および処刑対象の決定にプレイヤーごとの発話の傾向がどのように影響するのか, また, ゲーム全体のコミュニケーションの傾向とゲームの勝敗の関係について調査した. 分析の結果, 人間側, 人狼側の各プレイヤーが自陣営の勝利のために効果的なコミュニケーション戦略, および特定のコミュニケーションとプレイヤーの行動の関係が明らかとなった.

キーワード: 人狼ゲーム, 不完全情報ゲーム, 会話分析, 決定木

You Will Be Executed If You Keep Chatting Development of Tag Sets for Utterances in the Werewolf Game and Analysis of Action and Winning or Losing Using the Tags

MICHIMASA INABA^{1,a)} NAOMI OHATA^{1,†1} KENICHI TAKAHASHI¹ FUJIO TORIUMI²

Received: February 19, 2016, Accepted: September 6, 2016

Abstract: We focus on a communication game “Werewolf”. Our final objective is making a werewolf player agent. In this paper, as a first step, we analyze this game using players’ utterances. We define a tag set that describes player’s dialogue act and annotate utterances in the Werewolf BBS with these tags. By using these annotated utterances, we analyze the relationships between the trend of the types of the utterances and decision-making and victory or defeat. The results of the analysis clarify effective actions in werewolf in order to win.

Keywords: Werewolf Game, Imperfect Information Game, Dialogue Analysis, Decision Tree

1. はじめに

とある村に現れた人狼は人間と同じ姿をしており, 昼間には区別がつかないが, 夜になると村人たちを 1 人ずつ襲っていく. 疑心暗鬼になった村人たちは, 合議によって

人狼と思われる人物を 1 人ずつ処刑していくことにした……この設定のもと, プレイヤ間の対話により勝利を目指すのが人狼ゲームである. 村人は自分以外のプレイヤーに関する情報はいっさい与えられないため, 会話だけが唯一の手がかりとなる. 限られた情報の中で, 村人側はいかに人狼の正体を見破るか, 人狼側はいかに村人を欺き, 正体を隠し通すか, その駆け引きが人狼ゲームの醍醐味である. 本研究の目標は, この人狼ゲームを人と一緒にプレイできるエージェント (人狼知能) の実現にある. さらに, その人狼知能は単に強いだけでなく, 人に楽しさを感じさせ

¹ 広島市立大学
Hiroshima City University, Hiroshima 731-3194, Japan

² 東京大学
The University of Tokyo, Bunkyo, Tokyo 113-8654, Japan

^{†1} 現在, 株式会社ケイ・シー・シー
Presently with KCC Corporation

^{a)} inaba@hiroshima-cu.ac.jp

ることのできるものとするのが、我々の最終目標である。

近年、ゲームにおける人工知能の発展は著しく、1997年にチェスとオセロにおいてコンピュータが世界チャンピオンを破った。また、2013年に将棋において現役のプロ棋士を破り、2015年には情報処理学会のコンピュータ将棋プロジェクトがその役割を果たしたとして、終了宣言を行った。さらに2016年には囲碁においてGoogleの開発したAlphaGoがプロ相手に勝利した[1]。このようにボードゲームにおいてコンピュータが人間を圧倒するのに十分な計算資源と技術を獲得しつつある今、より一般的な要素を持つゲームにも対象を広げ、さらなる発展を目指す段階に来ている。

そこで我々はコミュニケーションにより勝敗が決定するゲーム「人狼ゲーム」に着目した。人狼ゲームの持つ一般性・普遍性の高さから、本研究が貢献できる領域はボードゲーム研究が主に対象としてきた探索手法や機械学習にとどまらない。エージェントインタラクション、言語処理、認知科学や心理学など含む領域が、人狼知能研究の対象とする領域である。人狼ゲームでは論理的な思考により相手の意図を読むことはもちろん、説得を行い、信頼を獲得することがきわめて重要である。これらは我々が社会活動において日常的に行っていることでもあり、様々な対話エージェントに適用可能な汎用コミュニケーション技術である。このような高度なコミュニケーション技術がゲームの勝利のためには必須であり、かつ勝敗による客観的な評価が可能である人狼ゲーム研究は、技術の発展を大きく後押しできる可能性がある。

そこで我々は、人狼ゲームをプレイできるエージェント(人狼知能)の研究を進めている。そのようなエージェントの構築には、ノンバーバル情報の理解・表出、他プレイヤーとの協調、自然言語処理など多数の課題が存在する。そこで、これまでに我々は人狼知能が会話を行うためのプロトコルの設計[2]や、そのプロトコルを用いてゲームを行うエージェントの構築キットの公開[3]、擬人化エージェントによる人狼ゲーム対戦システムの構築[4]などを行ってきた。これらの研究はコンピュータどうして人狼が対戦できる環境を整備し、競技会の開催を通じて集合知を集めることで、高度なコミュニケーションの実現のための課題解決を目指したものである。また、人間同士のプレイから人狼知能実現のための知見を得る試みとして、対面で行った人狼ゲームの分析も進めており、ノンバーバル情報がゲームに与える影響を調査した[5]。この研究では、対戦中の人間の仕草にアノテーションした動画の分析を行うことで、ゲームの勝敗に影響する仕草を明らかにした。

一方、本論文では対面で行った人狼ゲームではなく、Web上で行われたゲームログから勝利のために有効なコミュニケーション戦略を明らかにすることを目指す。そのような戦略が明らかとなれば、強い人狼知能を構築する際の指針

とすることができる。また、コミュニケーションとプレイヤーの行動の関係を分析することも本研究の目的である。人狼ゲームでは、プレイヤーの主観(他のプレイヤーに対する印象など)も行動に大きく影響する。そこで、プレイヤーの主観に大きく影響を与えるコミュニケーションと行動の関係を分析することで、客観的な情報だけでなく、主観的な情報も考慮できる人狼知能の構築に役立つ知見を得ることができる。これは、ただ強いだけではなく、人間らしい思考の実現や、主観的な情報を重視する度合いを変えて個性を出すなど、人が一緒にプレイして楽しい人狼知能の実現のために有用な知見である。

Web上で行う人狼ゲームはテキスト対話をベースにゲームが進行するため、対面で行うゲームにおいて重要なノンバーバル情報はいっさい勝敗に影響しないということが大きな特徴であり、また、議論の時間も対面の場合と比較して長いことから、コミュニケーションの内容にフォーカスした分析が可能である。しかし、現状では、人狼ゲームにおけるプレイヤーのコミュニケーションについての分析を行った研究は我々が知る限り存在しない。そこで、本論文では、ゲームの勝利のために有効なコミュニケーション戦略を明らかにするとともに、コミュニケーションとプレイヤーの行動の関係を分析するため、以下の2点を行う。

- 発話内容を表すタグの設計
- ゲーム内の発話へのタグ付与とタグを用いた分析

まず1点目であるが、実際のゲームデータの分析を行い、発話内容を意味や話題ごとに分類することで、発話の内容を示すタグを設計する。2点目は、設計したタグをゲームデータに付与することにより、自然言語による会話をマシンリーダブルな形に変換する。次にそのタグを用いてゲームの分析を行う。

分析およびタグ付与を行うデータは、オンラインネットゲームとして提供されている「人狼BBS*1」で行われたゲームのログを用いる。

2. 関連研究

Bravemanらによる研究[6]とYaoによる研究[7]では、ゲーム開始時のプレイヤーの数を n 、そのうちの人狼の数を m としたとき、人狼側が勝利する確率 $w(n, m)$ は m/\sqrt{n} に比例することを示した。その後、Migdalは人狼側が勝利する確率 $w(n, m)$ の厳密な式を示した[8]。

以上で述べた研究では、数理モデル化の際、単純化のため「占い師」など特殊能力を持つ役職は存在せず、村人と人狼のみでゲームが行われることを前提としている。しかし、実際には人狼BBSと同様、人狼と村人だけでなく、他の役職を含めてゲームが行われることは非常に多く、また、人狼と村人だけのゲームとはゲームの展開や性質が大きく

*1 <http://ninjinix.x0.com/wolf0/>

異なる。たとえば、人狼 BBS における各陣営の人数を、前述した Migdal の式にあてはめた場合（特殊な能力を持つ役職は「村人」として計算、 $n=15, m=3$ ）、人狼側の勝率は 97.1%と計算されるが、我々の調査によると人狼 BBS における実際の人狼側の勝率は 38.2%であり、大きく乖離している [9]。そのほかにも村人、占い師、狩人、人狼を含む 10 人村を対象とし、有効な戦略を分析した研究 [10] も存在するが、対話を行わないことを仮定するなど、いくつかの大きな制約を設けたうえで分析を行っている。本研究では、実際のゲームログを分析するため、近年頻繁に採用されているルールでプレイされた人狼ゲームの分析を行う点で他の研究とは異なる。

その他、人狼をプレイする人間の行動や心理的な側面に焦点を当て、様々な特徴を用いてプレイヤーの役職が人狼か否かの判定を行った研究もいくつか報告されている。たとえば、プレイヤーそれぞれの話の長さや回数、話を遮った回数などを特徴として用いた研究 [11] や、手や頭の動きを用いた研究 [12]、発話の長さを用いた研究 [13] などがある。しかし、以上にあげた研究において、ゲーム中の発話や議論の内容を対象としたものはない。

3. 人狼ゲーム

3.1 ルール

人狼ゲームは、人間側と人狼側に分かれ、自分の正体を隠しつつ他のプレイヤーの正体を探り、自陣営の勝利を目指すゲームである。プレイヤーはゲーム開始時に人間側か人狼側のどちらかにランダムに振り分けられる。その際、人間側のプレイヤーはどのプレイヤーがどちらの陣営に属しているかを知ることができない。一方、人狼側はどのプレイヤーがどの陣営に属しているかを知ることができる。

ゲームは昼と夜の 2 つのフェーズからなる。昼フェーズではすべてのプレイヤーが議論し、誰を人狼と疑い、処刑するかを投票により決定する。夜フェーズでは人狼側が人間側のプレイヤー 1 人を選び、襲撃する。処刑・襲撃されたプレイヤーはゲームから除外される。この昼フェーズと夜フェーズを繰り返し、最終的に人狼を全滅させた場合は人間の勝利となり、役職が人狼以外のプレイヤーの人数（狂人を含む）と人狼の人数が同数になった場合は人狼側陣営の勝利となる。

3.2 役職

ゲームには多くのバリエーションがあり、特殊な能力を持った役職をゲームに含めることもよく行われている。人狼 BBS の役職は、その中でもオーソドックスなものを中心に採用している。表 1 に人狼 BBS における役職とその能力、および所属する陣営を示した。

この中で狂人は、占い師と霊能者の能力では「人狼ではない」と判定されるが、人狼側に属し、人狼側の勝利が狂

表 1 人狼 BBS における役職
Table 1 Roles in the Werewolf BBS.

役職	陣営	特殊能力
村人	人間側	なし
占い師	人間側	夜フェーズにプレイヤー 1 人を指名し、そのプレイヤーが人狼か否かを知ることができる
狩人	人間側	夜フェーズにプレイヤー 1 人を指名し、そのプレイヤーを人狼の襲撃から守ることができる
霊能者	人間側	プレイヤーが死亡した際、そのプレイヤーが人狼か否かを知ることができる
共有者	人間側	ゲーム開始時にもう 1 人の共有者が誰であるかを知ることができる
人狼	人狼側	夜フェーズにプレイヤー 1 人を指名し、そのプレイヤーを襲撃できる（人狼全体で 1 人襲撃） 人狼しか聞こえない発話ができる。
狂人	人狼側	なし

人の勝利となる特殊な役職である。狂人は占い師や霊能者を騙って人間側を混乱させるのがセオリーである。

また、共有者は 2 人ペアで割り当てられる役職であり、ゲーム開始時にもう 1 人の共有者、すなわち人間側のプレイヤーを自分以外に 1 人だけ知ることができる役職である。共有者は 2 人同時に名乗り出ること、他のプレイヤーに自分たちが人狼ではないことを示すことが実際のゲームではよく行われる。これは、お互いに相手の役職を知っているプレイヤーは共有者以外には人狼しか存在しないため、共有者を騙って人狼が 2 人も名乗り出るとはきわめてリスクが大きく、ほとんど実行されないことを利用した戦略である。

3.3 人狼 BBS

人狼 BBS は Web 上で誰もが参加可能な Web ゲームとして公開されている。1 回のゲームは 9 人～15 人で行われ、ゲーム内の 1 日はリアルタイムの 1 日と同じ時間で進行する。役割ごとの構成人数はゲーム開始時のプレイヤー数によってあらかじめ決定されている。人狼 BBS では、何度かルールや役職ごとの構成人数の変更が行われており、本研究では同一のルール・人数・役職構成で行われたゲームデータを分析対象とする。なお、本論文で分析対象とするプレイヤー数 15 人（最大）の場合の役職構成は村人 6 人、人狼 3 人、占い師 1 人、狩人 1 人、霊能者 1 人、狂人 1 人、共有者 2 人である。

人狼 BBS には、他の人狼ゲームではほとんど採用されていない特殊なルールとして、発話回数制限と突然死がある。発話回数制限は、1 日に発言できる限度回数を超えて発言できないというルールである。突然死は、1 日に 1 度も発言しなかったプレイヤーは強制的に死亡するというルールで

ある。突然死が発生した場合、ゲームバランスが大きく崩れる場合があるため、本研究では突然死が発生したゲームは分析対象から除外する。

4. 人狼発話タグ

本研究では、人狼 BBS 中の発話に対して発話内容を表すタグを付与する。

発話を注釈付けするタグとしては、雑談の分析のために考案された Switchboard-DAMSL (Discourse Annotation and Markup System of Labeling) タグ [14] が知られている。Switchboard-DAMSL タグは雑談だけではなく、多人数による会議を収録した ICSI-MRDA (Meeting Recorder Dialog Act) コーパス [15] の発話注釈付けにも使用されており、人狼ゲーム内の発話にも付与可能である。しかし、このタグは「Statement-non-opinion (客観的事実)」や「Yes-No-Question (クローズドクエスション)」のように、発話の意味は扱わず、発話の表層的な側面をとらえるためのものである。人狼ゲームにおける戦略を分析するためには、ゲーム中の議論の内容を分析する必要があるため、発話の表層的な側面だけでは不十分である。また、同じく多人数による会議を収録した AMI Meeting コーパス [16] では、発話に対して独自のタグが付与されているが、こちらのタグも「情報提供 (Inform)」や「Suggest (提案)」などのように、発話の内容にまで踏み込んだタグではない。

そこで、本研究では人狼ゲームにおける発話の内容を表す人狼発話タグを新たに設計し、人狼 BBS の発話に付与する。タグの設計にあたり、人狼ゲームの議論を分析し、発話を内容ごとに分類することでヒューリスティックに個々のタグを定義する。その時点で分類できない場合は新たなタグを追加し、内容が重複するものは統合するといったことを繰り返すことで設計を行うが、その際、以下の基準により追加・統合を行うか否かを決定する。

(1) ゲームの特徴をとらえるタグ

自分の役職を明かすカムフラウト (coming-out) や他者を疑う発話 (suspicion) などは、人狼ゲームの戦略上不可欠な発話であることから個別のタグを定義する。

(2) 勝敗への影響の大小による細分化

予想・推測などにより疑いをかける suspicion と、役職固有の能力に基づいて他者の役職を宣言する decision については、意味が類似しているが、ゲームの勝敗に与える影響が大きく異なると考えられるため、タグを分割して定義する。

(3) 出現頻度

雑談を意味する idleTalk や自分が議論に参加できる時間を述べる schedule などは、勝敗への影響は小さくないと考えられるが、高頻度で出現するためタグを定義する。逆に、出現頻度が低いものはその他 (other) に

まとめる。

最終的に、20 種類のタグを定義した。次節では、各タグとその付与基準を例を用いて説明する。

4.1 タグの種類

coming-out

役職の表明 (CO: Coming-Out) に関する発話に付与する。自分の役職を表明する発話だけではなく、いつ役職を明らかにするかなどの発話にもこのタグを付与する。

例 1) 私は村人です

例 2) 占い師はいつ CO するのがいいですかね?

fortune-telling

占いに関する発話に付与する。次の占い先に関する発話にもこのタグを付与する。

例 1) では、占いについて話し合いましょう

例 2) 占い先は共有者の人に任せます

guard

狩人の護衛に関する発話に付与する。

例 1) 誰を守るといいと思いますか?

例 2) 狩人がいるなら、リーザを守ってほしい

attack

人狼の襲撃に関する発話に付与する。

例 1) 今日の襲撃は失敗してほしいですね

例 2) 次はヴァルターが襲われるんじゃないかな

execution

処刑に関する発話に付与する。

例 1) カタリナを吊るべきだ

例 2) 私が吊られても構いません

organizer

まとめ役に関する発話に付与する。まとめ役とは議論をリードし、結論を出すプレイヤーである。ルール上で定められた役職ではないが、人狼 BBS ではまとめ役を置いて議論を行うことが頻繁に行われている。

例 1) まとめ役は誰にしますか?

例 2) 共有者がまとめ役をやるべきだと思う

progress

現在ゲームがどのような状態であるかを述べた発話に付与する。

例) 現在、人狼が 3 人、人間が 5 人生き残っているという状態ですね

prediction

これからどのようなゲーム展開となるかを予想した発話に付与する。

例) 明日、村の勝利が決まらなければ、リーザと

ジムゾンを吊るして村人 2 人と人狼 1 人の最終決戦になるね

summary

参加者の意見をまとめる発話に付与する.

例) つまり, 今日 CO を求めるのが 3 人, 明日以降でもいいという人が 7 人ですね

problem

現在の状況に対して問題を提起する発話に付与する.

- 例 1) 今問題なのは, 人狼が全員生き残っているということだ
- 例 2) 占い師を全吊りすることで, 本当の占い師まで吊ることになってしまうぞ

consent

他者の意見・行動に同意する発話に付与する.

例) 私はカタリナの意見に賛成だ

suspicion

他者に疑いをかける発話に付与する. ただし, 人狼だけではなく, 占い師や霊能者と疑う発話にも付与する. また, 人狼側, もしくは人間側かという大まかな推理・予想に対してもこのタグを付与する.

- 例 1) ピーターは本物の霊能者っぽいね
- 例 2) リーザはすごく怪しい

reason

自身の発話や行動の理由を述べた発話に付与する.

- 例 1) というのも, 狩人はすでにいないと思うからです
- 例 2) 昨日発言しなかったのは, なかなか仕事が終わらなくて……

criticism

他者の意見や行動を批判する発話に付与する.

- 例 1) ヴァルターさん, 皆を混乱させるような発言は控えてください
- 例 2) モーリッツの意見は矛盾してるぞ

requestSpeaking

他者に発言を促す発話に付与する.

例) みなさん, 占い先について意見を出してください

decision

自分の能力に基づいて, 他者の役職を発表する発話に付与する. なお, 話者が真の能力者であるか否かは考慮せず付与する.

- 例 1) リーザは人間でした
- 例 2) もう 1 人の共有者はモーリッツです

correction

記入ミスや誤字脱字など, 単純なミスを訂正する発話に付与する.

例) さっきの発言は, モーリッツではなくリーザの間違いです.

schedule

議論に参加できる時間帯など, スケジュールに関する発話に付与する.

例) 私はだいたい夜 7 時ごろから顔を出せます

idleTalk

雑談など, ゲームに関係ない発話に付与する.

- 例 1) おはよう! 今から学校行ってきます.
- 例 2) ねえ, リーザ遊びに行こうよ!

other

上記のタグすべてに該当せず, ゲームに関する発話に付与する.

- 例 1) 序盤では白確定を増やすことでグレーの数を減らすべきだと思う.
- 例 2) 不審な行動をとったせいで怪しまれるのしょうがない.

4.2 タグ付与の手順

人狼 BBS の各発話 (BBS 上の 1 回の投稿) は複数の文からなっていることも多い. タグ付与はまず文ごとに行い, 各文に対して少なくとも 1 個以上の仮タグを付与する. そして, 1 発話中の全文に仮タグを付与した後, 重複している仮タグを 1 つにまとめ, それを最終的なタグ付与結果とする. たとえば「私は占いはリーザで, 吊りはトーマスがいいと思う. リーザは人間だと思うから. トーマスは人狼な気がするから」という発話に対しては, 1 文目は「fortune-telling」と「execution」, 2 文目は「suspicion」と「reason」, 3 文目は「suspicion」と「reason」が仮タグとして付与される. 重複した仮タグは 1 つにまとめられるため, 最終的に付与されるタグは「fortune-telling」, 「execution」, 「suspicion」, 「reason」の 4 つとなる. 以降, 本論文におけるタグとは, この最終的に付与されたタグのことを指すものとする.

タグ付与対象のゲームは, 参加者が最大の 15 人で突然死が発生していないものから, 人狼側勝利のゲーム数と人間側勝利のゲーム数が同数になるようにランダムに選択した. なお, 本研究で収集したデータにおいて最も多く行われているのが 15 人のゲームである*2. タグ付与は本論文の著者の大学生 1 人が行った.

人狼 BBS における投稿には, 全員に公開される通常発話のほか, 人狼のみが投稿・閲覧できる「ささやき」, 生存

*2 1,598 ゲーム中, 1,512 ゲーム (94.6%) が 15 人のゲーム.

表 2 タグを付与したデータの例 (人狼 BBS 166 僻地の村 3 日目の議論の一部)

Table 2 Example of tagged data.

名前	発話	タグ
ヤコブ	最後に、まだ潜伏している真の占い師がいるならあえて言おう。もう CO しないでくれ。ここででてくることは君が真の占い師であったとしてもさらに事態を混乱させる危険性のほうが大きくなる。そのために今回は CO はしないでほしい。これが俺の考えだ。	other coming-out reason
カタリナ	すぐに出かけますが、一旦戻ってきました。結果をお伝えするのを忘れていたです。トーマスさんは、人狼でした。そしてなぜ神父様が・・・	idleTalk decision attack
カタリナ	今は時間がないので、また夜帰ってきてから発言します。あと、変なタイミングでの CO、ほんとにごめんなさい。	schedule coming-out
ペーター	カタリナさんの他に霊能者 CO が在るかどうかがわからないと、ヤコブさんの言いたいことがよくわからないので、全員が発言するまで、待っててもいいかな？ カタリナさん以外に霊能者さんが CO しなかったら、ヤコブさん、今日の吊りはあなたが吊られることで信用を得て、カタリナ黒+ニコラス黒を証明してください。私が今言えることはそれだけです。	coming-out other reason execution
レジーナ	詳しいことは夜になってみんなが集まってから議論しようじゃないか。 今夜は冷えるみたいだからビーフストロガノフを作ったよ。 オットーのパンと一緒に食べると最高に美味しいから試してみてください。ワインも用意しておくよ。	idleTalk

表 3 タグ付与結果

Table 3 Annotation result.

ゲーム数	24
日数	151
総発話数	19,263
総タグ数	36,571

しているプレイヤーのみが投稿でき、ゲーム終了後に公開される「独り言」、死亡しているプレイヤーのみが投稿・閲覧できる「うめき」がある。このうち、タグ付与を行ったのは通常発話のみであり、それ以外の投稿は本論文では扱わない。また、ゲーム初日*3とゲーム終了後の1日はプレイヤー同士の交流の場として設定されており、ゲームに関係する会話はなされないためタグ付与は行わない。

実際のタグ付与の例を表 2 に示す。また、タグ付与の結果を表 3 に、最終的に付与した各タグの出現割合を順に並べたものを表 4 に示す。なお、陣営別のタグの出現割合についても確認を行ったが、人間側・人狼側ともに表 4 とほぼ同様の順位・出現割合であり、大きな差は確認できなかった。

4.3 人狼発話タグの妥当性

本節では定義した人狼発話タグが、タグ付与者が替わっても一貫した付与が可能な設計であるかを調査する。そのために、タグ付与を行った大学生（以下、アノテータ A とする）とは別のアノテータ B がタグ付与を行い、2 人同士のタグの一致率を調査した。アノテータ B は人狼ゲームのルールを理解しているが、対話へのアノテートや対話研究はこれまでにない大学生である。一致率の調査のため、アノテータ A がタグ付与を行った 151 日分の

*3 人狼 BBS 内ではプロローグと表記。

表 4 各タグが付与タグ全体に占める割合

Table 4 Ratio of each tag.

順位	タグ	割合
1	idleTalk	20.22%
2	other	16.00%
3	reason	14.69%
4	execution	9.23%
5	fortune-telling	8.02%
6	suspicion	6.81%
7	coming-out	5.89%
8	requestSpeaking	4.64%
9	consent	2.64%
10	schedule	2.23%
11	criticism	1.81%
12	prediction	1.71%
13	progress	1.59%
14	organizer	1.12%
15	correction	0.79%
16	decision	0.76%
17	summary	0.74%
18	attack	0.44%
19	guard	0.43%
20	problem	0.24%

データから 3 日分のデータをランダムで抽出し、改めてアノテータ B がタグ付与を行った。対象となったデータは同一ゲーム中の連続した 3 日分のデータではなく、それぞれ異なるゲームから抽出された。また、タグ付与対象の発話数は 472 個であった。

判定者間の一致度を測るため、Cohen の κ による評価を行う。ただし、Cohen の κ は、1 つの判定対象に複数の属性を付与する今回のような場合には直接適用できないため、特定のタグに着目し、各発話に対してそのタグを

表 5 タグ付与妥当性評価

Table 5 Evaluation of the appropriateness of the annotation.

A による付与タグ数	1,018
B による付与タグ数	1,212
一致タグ数	738
一致率	64.0%

「付与する」か「付与しない」の2種類を判定したものとして計算した。その結果、20種類のタグにおけるCohenの κ の平均は0.67(分散は0.05)であり、かなりの一致を示した。個別に見てみると、idleTalk、およびotherの一致度が低くなっていた。idleTalkのCohenの κ は0.55であった。これは雑談か、あるいはゲームに関係あるものなのかの分類が難しい発話が多く存在していることが原因である。また、otherはタグのカバーする範囲が大きいため、アノテータによる判断が難しくなり、Cohenの κ は0.47という結果となった。しかし、いずれも中程度以上の一致は確認できた。一方、内容が明確に発話されることの多いexecution、fortune-telling、coming-outは一致度が高く、それぞれCohenの κ は0.82、0.84、0.78と、ほぼ一致を確認できた。したがって、本論文で提案したタグは、アノテータが異なっても一貫したタグ付与が可能な設計であると確認できた。

さらに、雑談中の発話にタグ付与を行った文献[17]で用いられた以下の式で示される指標による一致率も計算した。

$$\text{一致率} = \frac{(\text{一致したタグ数}) \times 2}{\text{AとBが付与したタグの総数}} \times 100$$

この式により、1つの判定対象に複数の属性を付与する場合の一致率が計算できる。

結果を表5に示す。文献[17]では、「客観的事実」などの発話行為を意味するタグであるDialogue Actタグを47種類、「前発言の詳細化」などの発話間の関係を意味するタグであるRhetorical Relationタグを16種類それぞれ定義している。そこでの一致率は、Dialogue Actタグでは65.5%、Rhetorical Relationタグでは59.6%であった。表に示したとおり、本研究で定義したタグの一致率は64.0%であり、文献[17]におけるタグと比較すると、同程度に一貫したタグ付与が可能なタグ設計になっていることが確認できた。

タグの信頼性を確保するため、アノテーションは複数人で行い、その合議によってタグを決定する場合もあるが、本研究では、分析に用いるタグを付与したアノテータは1人である。しかしながら、比較的高いCohenの κ 、および一致率が確認できたことから、アノテータが1人であっても、付与されたタグの信頼性は確保されていると考えられる。

5. 人狼ゲームの分析

本章では、タグ付与を行った人狼BBSのデータを用い

て2種類の分析を行い、人狼知能構築のために有用な知識を獲得する。1つ目はプレイヤー別の発話傾向の着目した分析である。発話に付与されたタグをプレイヤー別に集計し、どのような発話が処刑対象と襲撃対象の選択に影響を与えるかを分析する。もう1つはゲーム全体の発話傾向に着目した分析である。タグをゲーム単位で集計し、タグの出現傾向と勝敗の関係を分析する。ここでは、人狼側と人間側の発話を個別に分析を行うことで、各陣営ごとの有効な戦略を明らかにする。

5.1 発話傾向に基づく分析

本節では、発話に付与されたタグをプレイヤー別に集計し、どのような発話が処刑対象と襲撃対象の選択に影響を与えるかを分析する。そこで、ゲーム内の各日ごとの対象となったプレイヤーと、それ以外のプレイヤーの発話に付与されたタグを別々に集計し、各タグが全タグに占める割合の大小を2者間で比較する(対象以外のプレイヤーのタグはまとめて集計する)。なお、割合の比較は1ゲーム単位ではなく、1日単位で行う。こうすることで、対象となったプレイヤーとなっていないプレイヤーの発言傾向の差を分析することが可能となる。ただし、分析対象とするタグは、タグ付与を行った全データにおける出現割合が5.00%を超えたタグとし、それ以下のタグは分析には使用しない。これは、出現割合が小さいタグは、1日に1度もそのタグが出現しないデータも多く、プレイヤー間の比較が困難なためである。加えて、その他を意味する「other」タグは分析には使用しない。これは、otherタグの付与される発話の内容は多種多様であるため、仮にotherタグが重要であることが判明しても、エージェントの構築に応用することは困難なためである。また、1日目には処刑対象と襲撃対象の選択が行われないことから、1日目のデータは使用しない。発話回数が少ない場合、タグの数が少なくても割合が大きくなってしまふことから、1日の発話が10回以下であったプレイヤーが対象となった日のデータも分析には使用しない。

5.1.1 処刑対象となったプレイヤーの発話傾向分析

図1に処刑対象プレイヤーとそれ以外のプレイヤーのタグごとの出現割合を各日ごとに比較し、処刑対象プレイヤーの方が割合が大きかった日数(データ数)をタグ別に示した。分析対象となったデータは94日分であった。図から、idleTalkは52日(データ)となっていることが分かるが、これは処刑対象となったプレイヤーの発話に付与されたタグにおけるidleTalkタグの占める割合と、その処刑対象プレイヤー以外のすべてのプレイヤーの発話に付与されたタグにおけるidleTalkタグの割合の2つを比較したとき、処刑対象となったプレイヤーの方が割合が大きかったデータが52個存在したことを意味する。また、表中の「**」は、94日分のデータをまとめ、各タグの出現割合を処刑対象プレイヤーとそれ以外のプレイヤーの間で比率の差の検定を実施した結

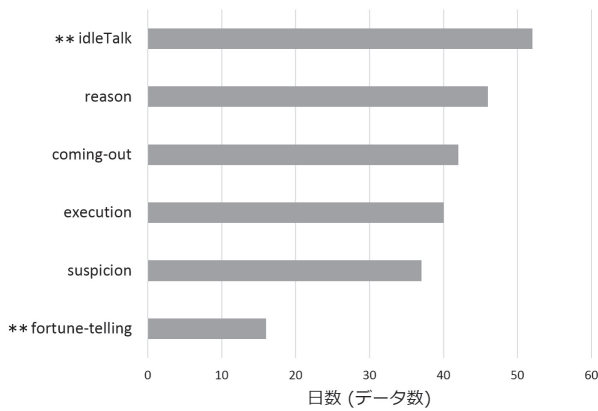


図 1 処刑対象プレイヤーの発話傾向

Fig. 1 Utterance tendency of executed players.

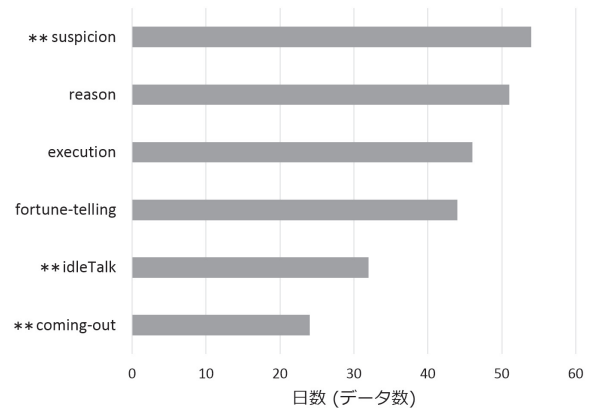


図 2 襲撃対象プレイヤーの発話傾向

Fig. 2 Utterance tendency of attacked players.

果である。「**」は有意水準 1% で有意な差が認められたことを示す。今回の結果では、idleTalk と fortune-telling について、有意水準 1% で優位な差が認められた。

ここから、処刑対象になったプレイヤーは、雑談 (idleTalk) が多いことが分かる。雑談ばかりしているプレイヤーは、他のプレイヤーに議論を行ううえで役に立たないという印象を与える場合があると考えられる。特に、人狼の正体につながるような情報がほとんどないような状況では、処刑対象の選択はそういった印象の良し悪しによって決定されることも多い。また、雑談によって話を逸らせようとしていると疑われる場合もあると考えられる。実際のデータを見ても、「発言にあまり内容がない」や「雑談で発言数をかせいでるようにしか見えない」といったことを処刑対象を選んだ理由としてあげたプレイヤーが存在していた。一方、処刑対象になったプレイヤーは、占い (fortune-telling) に関する発話が少ないことも分かる。これは、idleTalk とは逆に、重要な話題である占いに関する議論をしっかり行っているプレイヤーは、他のプレイヤーから良い印象を持たれ、処刑対象になりにくくなる傾向があると考えられる。

以上の分析から、処刑対象にならないためには雑談をせず、占いなどの重要な話題についても議論に参加することが重要であると考えられる。当然のことながら、処刑対象になる理由は雑談だけが理由ではなく、占い師によって人狼と見破られた場合や、言っていることにつじつまが合わないなど様々な理由が考えられる。しかし、ゲーム序盤は手がかりが少ないため、他のプレイヤーからの心証の悪化は致命的となる可能性があることから、雑談はできるだけ控えるべきであるといえる。

5.1.2 襲撃対象となったプレイヤーの発話傾向分析

図 2 に襲撃対象プレイヤーとそれ以外のプレイヤーのタグごとの出現割合を比較し、襲撃対象プレイヤーの方が大きかった日数をタグ別に示した。分析対象となったデータは 102 日分であった。また、suspicion, idleTalk, coming-out について、有意水準 1% で優位な差が認められた。ここから、襲

撃対象となったプレイヤーは、疑いをかける発話 (suspicion) が多いことが分かる。一方で、襲撃対象となったプレイヤーは、雑談 (idleTalk) が少ない。これは、疑いに関する発話 (suspicion) とは逆に、雑談は人狼にとって不利にならないため、襲撃対象から外れやすいと考えられる。また、襲撃対象となったプレイヤーは、役職の表明に関する発話 (coming-out) も少ない。これは、占い師と霊能者が自分の役職を表明する際、ほとんどの場合偽の占い師・霊能者も役職を表明することから、どちらが本物の能力者かを見極めるため (片方が狂人の場合)、偽の能力者を騙っている人狼の正体を隠すため、能力者を守る狩人の襲撃を優先するためなどの理由により、役職を表明したプレイヤーの襲撃を見送るケースが多いことがその一因であると考えられる*4。

以上の分析から、襲撃対象とならないためには、疑いをかける発言を控えることが重要であると考えられる。しかし、人狼ゲームにおいて他のプレイヤーに疑いをかけることは戦略上不可欠である。したがって、占い師や霊能者など、特殊な能力を持つプレイヤーは疑いをかける発言をできるだけ控えて襲撃を回避し、逆に能力を持たない村人が必要に応じて疑いをかけ、人狼を探し出していくことが人間側勝利のために有効であると思われる。

5.2 決定木による分析

前節では、プレイヤーごとにタグの集計を行ったが、本節では、タグをゲーム単位で集計し、タグの出現傾向と勝敗の関係性を分析する。分析には、1 ゲーム中に出現した各タグの全タグに占める割合を属性、勝利陣営をクラスとし、これを学習データとして構築した決定木を用いる。決定木を用いるのは可視化による分析の容易さを重視するためである。使用するタグは「other」以外で出現割合が 5.00% を超えた 6 種類とする。また、決定木の構築には

*4 最終的には占い師や霊能者は襲撃されることが多いが、今回はタグの出現した当日の襲撃対象プレイヤーを分析しているため、過去の日付で役職を表明したかということは考慮していない。

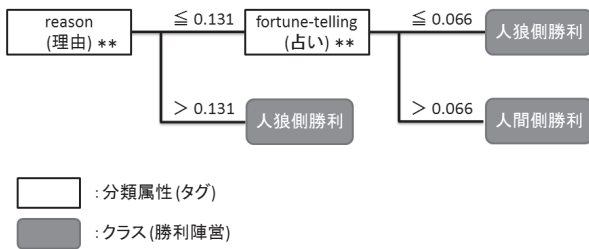


図 3 人間側プレイヤーの発話から構築した決定木

Fig. 3 Decision tree constructed using human-side players' utterances.

Weka [18] を用いた。決定木学習アルゴリズムは J48 を使用し、leave-one-out 交差検証により分類精度を求めた。

分析は、人間側（村人、占い師、狩人、霊能者、共有者）の発話に付与されたタグのみを用いた場合、人狼側（人狼、狂人）の発話に付与されたタグのみを用いた場合、およびゲーム全体のタグを用いた場合の3種類について行った。

5.2.1 人間側プレイヤーの発話傾向と勝敗の分析

構築した決定木を図 3 に示す。分類精度は 79.2%であった。

決定木より、理由（reason）が多ければ人狼側の勝利となっていることが分かる。また、理由が少ない場合、占いに関する発話（fortune-telling）が少ないと人狼側の、多いと人間側の勝利となっている。なお、図中の「*」は対応するタグの割合だけを用いて、人狼側勝利 12 ゲームと人間側勝利 12 ゲームの間で Welch の t 検定を行った結果（以下、「t 検定の結果」とする）である。「**」は有意水準 1%、「*」は有意水準 5%で有意な差が認められたことを示す。今回の場合、人狼側が勝利したゲームと人間側が勝利したゲームにおける人間側プレイヤーの reason タグの割合には、有意水準 1%で有意な差が認められた。つまり、reason タグが多いと人狼側が勝利するということが統計的に有意であるということの意味する。同様に、fortune-telling タグについても有意水準 1%で有意な差が認められた。

理由を述べることは、自分の考えを公開するというものである。したがって、理由が多いことは、各プレイヤーの考え方やそれに基づく行動を予測するための情報がより多くなることを意味するため、人狼側にとっては戦略を練るのが容易になっていると考えられる。したがって、人間側プレイヤーは必要以上に自分の考えを発言しないことが勝利につながると考えられる。さらに、昼フェーズで重要な話題は占いにすることであると分かる。したがって、人間側は占いに関する話題を積極的に議論することが勝利のために重要である。一方で、人狼側は占いに関する話題についてあまり議論させないように話を誘導していくことが重要である。

5.2.2 人狼側プレイヤーの発話傾向と勝敗の分析

構築した決定木を図 4 に示す。分類精度は 87.5%であっ

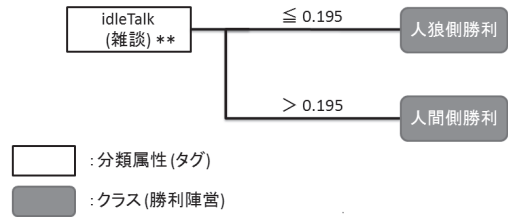


図 4 人狼側プレイヤーの発話から構築した決定木

Fig. 4 Decision tree constructed using werewolf-side players' utterances.

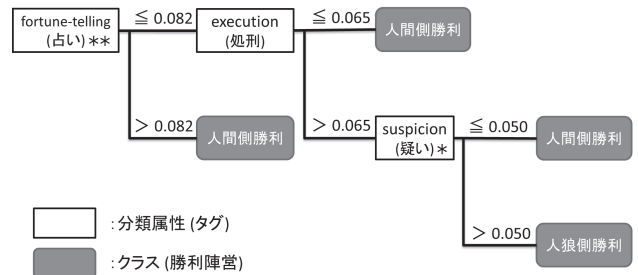


図 5 人狼側プレイヤーの発話から構築した決定木 (idleTalk 除外)

Fig. 5 Decision tree constructed using werewolf-side players' utterances (excluding idleTalk).

た。この決定木は、雑談（idleTalk）が少なければ人狼側勝利、多ければ人間側勝利となっている。また、idleTalk タグのみを用いた t 検定の結果、有意水準 1%で有意な差が認められた。この結果は、5.1.1 項で述べた雑談が多いと処刑対象になるという結果と一致している。つまり、人狼側プレイヤーの雑談の多さはその処刑につながり、ひいては人狼側の敗北につながるということである。特に、人狼の死因は処刑のみであるため、役職が人狼のプレイヤーは雑談を控えることが重要であるといえる。

ここで、idleTalk 以外のタグに関しても分析を行うため、idleTalk タグを除外して決定木の構築を行った。構築した決定木を図 5 に示す。この決定木の分類精度は 75.0%であった。

決定木より、占いに関する発話（fortune-telling）が多ければ人間側が勝利していることが分かる。また、処刑に関する発話（execution）と疑いをかける発話（suspicion）が少ない場合にも人間側が勝利している。t 検定の結果、fortune-telling タグは有意水準 1%で、suspicion タグは 5%で有意な差が認められた。execution タグに関しては有意な差は認められなかった。この結果は 5.2.1 項で述べたように、人間側にとって占いに関する議論が重要であることを示している。一方、占いに関する発話が少ない場合、処刑に関する発話と疑いをかける発話が多いと人狼側の勝利となっている。execution タグに関しては有意な差は認められなかったことをふまえると、人狼側にとって有効な戦略は、積極的に他のプレイヤーに疑いをかけていくことであることが示唆された。

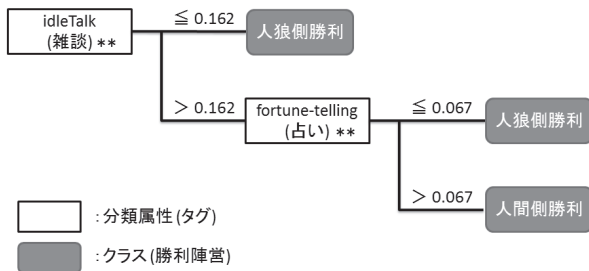


図 6 全プレイヤーの発話から構築した決定木

Fig. 6 Decision tree constructed using all players' utterances.

5.2.3 全プレイヤーの発話傾向と勝敗の分析

構築した決定木を図 6 に示す。分類精度は 91.7%であった。

決定木より、雑談 (idleTalk) が少ない場合に人狼側が勝利となり、雑談が多い場合には、占いに関する発話 (fortune-telling) が多く人間側、少ないと人狼側の勝利となっている。t 検定の結果、idleTalk タグ、および fortune-telling タグはそれぞれ有意水準 1% で有意な差が認められた。この結果は、人間側プレイヤーのみの場合と人狼側プレイヤーのみの場合の結果を合わせた結果となっている。すなわち、図 4 では、雑談が少ないと人狼側、多いと人間側が勝利となり、図 3 と図 5 ではともに、占いに関する発話が多いと人間側が勝利している。ここから読み取れることは、人狼側にとっては雑談の少なさが、人間側にとっては占いに関する議論が重要だということである。村人以外の役職、特に人狼は雑談をするべきではない。人狼の死因は処刑のみであり、処刑対象から逃れることが勝利への近道だからである。また、人間側は占いに関する議論を綿密に行うことが勝利のために重要である。

5.3 分析のまとめ

以上の分析により判明した、各陣営のプレイヤーが自陣営の勝利のために効果的な行動を以下にまとめる。

- 人間側プレイヤー
 - 占いに関することを最優先で話し合う (5.2.1 項より)
 - 能力者は雑談しない (5.1.1 項より)
 - 能力者は疑いをかけるのを控える (5.1.2 項より)
 - 必要以上に自分の考えを明らかにしない (5.2.1 項より)
- 人狼側プレイヤー
 - 雑談はしない (5.1.1, 5.2.2 項より)
 - 占いに関して議論させないように話を誘導する (5.2.1 項より)
 - 積極的に他のプレイヤーに疑いをかける (5.2.2 項より)

6. おわりに

本論文では、人狼 BBS のゲームデータを分析し、人狼ゲームにおけるプレイヤーの発話内容を表現するタグを設計

した。次に、設計したタグを人狼 BBS におけるプレイヤーの発話に付与し分析を行った。分析では、襲撃対象、および処刑対象の決定にプレイヤーごとの発話の傾向がどのように影響するのかを調査した。また、ゲーム全体の発話の傾向とゲームの勝敗の関係について分析した。分析の結果、人間側、人狼側の各プレイヤーが自陣営の勝利のために効果的なコミュニケーション戦略、および特定のコミュニケーションとプレイヤーの行動の関係が明らかとなった。

しかし、人狼知能の実現にはまだ課題も多い。その 1 つとして、人狼における戦略の定式化があげられる。本研究では、各タグの頻度に基づいた分析を行ったが、議論のプロセスに踏み込んだ分析は行っていない。人狼には人間側、人狼側それぞれに一定の戦略が存在している。議論のプロセスを分析することにより、そういった戦略がゲームの勝敗にどのような影響を与えているのかを明らかにし、プレイヤー間の会話から戦略を抽出する技術を確認することが人狼知能設計のための大きな課題である。その際、本研究で定義した人狼発話タグは有用であると思われる。

さらに、本研究では、タグを手で付与したが、アノテーションのコストが非常に大きいことから、タグの自動付与も今後の課題である。しかし、単純な方法では自動付与が難しいことがこれまでに判明しており、発話に含まれる単語を素性とし、Support Vector Machine を用いて各タグの付与の有無で 2 値分類したところ、最も性能が良かった idleTalk タグでも精度 0.55、再現率 0.68、F 値 0.60 にとどまった。ただし我々は、CO タグ、decision タグに関しては F 値がそれぞれ 0.98、0.78 で自動付与可能な手法 [19] を提案しており、今後はそれ以外のタグに関しても高い精度で付与可能な手法を検討していく。

謝辞 人狼 BBS のデータ使用を許可していただいた ninjin 氏に感謝いたします。なお、本研究は JSPS 科研費 15K12180 の助成を受けたものです。

参考文献

- [1] Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., van den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T. and Hassabis, D.: Mastering the game of Go with deep neural networks and tree search, *Nature*, Vol.529, pp.484-503 (2016).
- [2] 大澤博隆, 鳥海不二夫, 片上大輔, 篠田孝祐, 稲葉通将: 人狼ゲームのプロトコル設計: 推理と説得のプロトコル, 第 24 回インテリジェント・システム・シンポジウム (FAN2014) (2014).
- [3] 鳥海不二夫, 梶原健吾, 大澤博隆, 稲葉通将, 片上大輔, 篠田孝祐: 人狼知能プラットフォームの開発, デジタルゲーム学会 2014 年度年次大会 (2015).
- [4] 片上大輔, 小林 優, 大澤博隆, 稲葉通将, 篠田孝祐, 鳥海不二夫: 擬人化エージェントを用いた人狼対戦システムの開発, ゲームプログラミングワークショップ 2014 論文

- 集, Vol.2014, pp.133-134 (2014).
- [5] 高久奨乃, 稲葉通将, 大澤博隆, 篠田孝祐, 西野順二, 鳥海不二夫, 片上大輔: 人狼ゲームにおいてノンバーバル情報が与える影響について, 第2回 HSS デザインコンテスト 2014 (2014).
- [6] Braverman, M., Etesami, O. and Mossel, E.: Mafia: A theoretical study of players and coalitions in a partial information environment, *The Annals of Applied Probability*, pp.825-846 (2008).
- [7] Yao, E.: A Theoretical Study of Mafia Games, Arxiv preprint arXiv:0804.0071 (2008).
- [8] Migdal, P.: A mathematical model of the Mafia game, Arxiv preprint arXiv:1009.1031 (2010).
- [9] 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, Vol.2012, No.6, pp.144-147 (2012).
- [10] 畢 晓恒, 田中哲朗: 対話のない人狼ゲームの戦略, ゲームプログラミングワークショップ 2015 論文集, pp.25-30 (2015).
- [11] Chittaranjan, G. and Hung, H.: Are you awerewolf? detecting deceptive roles and outcomes in a conversational role-playing game, *IEEE Intence Conference on Acoustics Speech and Signal Processing (ICASSP)*, pp.5334-5337 (2010).
- [12] Xia, F., Wang, H. and Huang, J.: Deception Detection Via Blob Motion Pattern Analysis, *Affective Computing and Intelligent Interaction*, pp.727-728 (2007).
- [13] Zhou, L. and Sung, Y.: Cues to deception in online Chinese groups, *Proc. 41st Annual Hawaii International Conference on System Sciences*, pp.146-146 (2008).
- [14] Jurafsky, D., Shriberg, E. and Biasca, D.: Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, University of Colorado, Institute of Cognitive Science Technical Report, pp.1-61 (1997).
- [15] Shriberg, E., Dhillon, R., Bhagat, S.V., Ang, J. and Carvey, H.: *The ICSI Meeting Recorder Dialog Act (MRDA) Corpus*, Defense Technical Information Center (2004).
- [16] Carletta, J., Ashby, S., Bourban, S., Flynn, M., Guillemot, M., Hain, T., Kadlec, J., Karaikos, V., Kraaij, W., Kronenthal, M., et al.: The AMI meeting corpus: A pre-announcement, *Machine Learning for Multimodal Interaction*, pp.28-39, Springer (2006).
- [17] 徳久良子, 寺寫立太: 雑談における発話のやりとりと盛り上がりとの関連, 人工知能学会論文誌, Vol.21, No.2, pp.133-142 (2006).
- [18] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. and Witten, I.H.: The WEKA data mining software: An update, *ACM SIGKDD Explorations Newsletter*, Vol.11, No.1, pp.10-18 (2009).
- [19] 平田佑也, 稲葉通将, 高橋健一, 鳥海不二夫, 大澤博隆, 片上大輔, 篠田孝祐: プレイログから獲得した行動選択確率を用いた人狼ゲームのシミュレーション, 第29回人工知能学会全国大会 (2015).



稲葉 通将 (正会員)

平成 24 年名古屋大学大学院情報科学研究科博士後期課程修了, 同年広島市立大学大学院情報科学研究科助教, 現在に至る. 対話システム, 対話処理に関する研究に従事. 博士 (情報科学), 電子情報通信学会, 人工知能学会, 言語処理学会各会員.

語処理学会各会員.



大畠 菜央実

平成 23 年広島市立大学情報科学部知能工学科卒業. 同年株式会社ケイ・シー・シー入社. 在学中は主に対話処理の研究に従事.



高橋 健一

昭和 52 年名古屋工業大学情報工学科卒業. 昭和 54 年同大学大学院工学研究科修士課程修了. 同年同大学工学部助手. 同大学講師, 助教授を経て, 平成 6 年広島市立大学情報科学部教授, 現在に至る. 工学博士. 主に機械学

習, パターン情報処理の研究に従事. 電子情報通信学会, IEEE 各会員.



鳥海 不二夫

平成 12 年東京工業大学大学院理工学研究科機械制御システム工学専攻博士課程修了, 同年名古屋大学情報科学研究科助手, 平成 19 年同助教, 平成 24 年東京大学大学院工学系研究科准教授, 現在に至る. エージェントベースシミュレーション, 人工市場, ソーシャルメディア等の

研究に従事.