

人狼知能大会における自然言語処理部門の構想

狩野 芳伸

静岡大学 情報学部 行動情報学科
kano@inf.shizuoka.ac.jp

稲葉 通将

広島市立大学 大学院情報科学研究科 知能工学専攻
inaba@hiroshima-cu.ac.jp

1 はじめに

近年、対話システムの利用が注目を集めている。従来多く見られたタスク指向型対話システムに加え、いわゆる雑談的な対話を行う非タスク指向型対話システムも多くみられる。対話システムが必要とする要素はテキスト入出力に限ったとしても、語彙、構文から意味、論理まで、完成度を上げていくと最終的にはほとんどあらゆる知的要素が必要となろう。

人狼ゲームは基本的に制約のない会話を通じてのみ行うゲームであり、対話システムに必要な要素を包含している。一方で人狼ゲームを題材にすることは、ゲームの勝敗という目的があるために状況を限定しつつ一步一步研究を進めようという利点がある。本稿では、対話システムの発展に寄与する研究の切り口として、会話ゲーム「人狼」のプレイヤー自動化を試みる人狼知能大会における、自然言語を用いる部門の構想について述べる。

2 人狼ゲーム

人狼ゲームの設定

人狼ゲームは 10 人程度のプレイヤーで行うゲームである。各プレイヤーは村人陣営と人狼陣営に分かれてそれぞれの陣営の勝利を目指す。ゲーム開始前に所属陣営を決定するが、基本的に村人には他プレイヤーの所属陣営は公開されない。村人陣営の目的は人狼をすべて見つけ出し追放（ゲームから排除）することで、人狼陣営の目的は村人を食べ尽くす（ゲームから排除）ことである。

人狼陣営のプレイヤーは排除されないよう、村人陣営の振りをするようになる。プレイヤーは基本的に会話のみを通じてお互いの正体を探るが、何の情報も無い状態では人狼側が有利になりすぎるため、村人側に有利な特殊能力を持ったいくつかの「役職」が主に村人側に用意されている。ゲームの設定により異なるが、一日一人だけ指定したプレイヤーが「人狼かどうか」を知ることができる（ただし本人以外にはそれが真かどうか、占ったかどうかも知らされない）「占い師」、人狼陣営だが自身は人狼が誰だかを知らず、能力も持たないため、場を混

乱させることによって人狼陣営に貢献することの多い「裏切者」などがある。

人狼ゲームの流れ

ゲームは「一日」を基準とするターンを繰り返す。一日の前半を「昼」と呼び、プレイヤーは決められた時間、自由に会話を行う。後半を「夜」と呼び、ゲームから排除するプレイヤーを投票によって決定する。村人陣営は人狼陣営のプレイヤーを排除しようとし、人狼陣営のプレイヤーは村人陣営のプレイヤーを排除しようとするだろうが、表面上はあくまで「人狼と思われる人」を排除するための投票になる。最大得票を得たプレイヤーはゲームから排除（「追放」）される。

次に、全プレイヤーが顔を伏せた状態で、司会者が様々な役職のプレイヤーに声をかけ、声をかけられたプレイヤーはジェスチャーなどで他のプレイヤーにわからないよう、特殊能力に従った行動をとる。例えば、人狼はゲームから排除するプレイヤーを一人決定する（「襲撃」）。

ここまでが「一日」のターンで、これをどちらかの陣営が勝利条件を満たすまで繰り返す。プレイヤーの勝敗は所属陣営の勝敗できまるので、途中で排除されたかどうかはプレイヤーの勝敗に直接関係がない。村人陣営の勝利条件は「すべての人狼を排除すること」、人狼陣営の勝利条件は「人狼以外のプレイヤーの数を人狼の数と同数またはそれ未満にする」ことである。

人狼ゲームは元々ロシアで遊ばれていたゲームでマフィアと呼ばれていた。それがアメリカで商品化され、日本でも「汝は人狼なりや」「タブラの狼」などいくつかのゲームセットが発売されている。現在では、インターネット越しに掲示板形式で人狼ゲームを長期にわたって行う BBS 人狼、短期間で行う短期人狼などがある。人狼をプレイする人口は現在増え続けているが、上の年代になるにつれて知名度は減少しており、若者のゲームであると言える。

3 人狼知能プロジェクトと大会

人狼知能プロジェクト

人狼知能プロジェクト[1][2][3]では、人狼知能プロトコルと対戦用の人狼知能サーバの情報

を公開することで、多く研究者や開発者がエージェントの作成に参加できる環境を整えてきた。

将棋やチェスなど盤面上にすべての情報が開示される完全情報ゲームとは異なり、人狼は不完全情報ゲームである。この特徴と、人狼が基本的に会話を通じてのみ行われるということが、プレイヤーの自動化という視点でとらえたときに特有の興味深い研究テーマを生み出している。

まず、状況判断がプレイヤーの言動に依存するため、おおくの局面で本質的に「正解」を知りえず、「推理」が必要である。推理する際にも、他プレイヤーの意図をモデル化するという高度な作業が必要となる。こうした推理やモデルの上に、人狼陣営はいやおうなく「嘘をついて騙す」ことが求められる一方、村人陣営はそえを「見破る」必要がある。これは見方を変えると、いかに他者を「説得」し「信頼を得るか」ということでもある。

さらに、実際のゲームでは表情や音声、ジェスチャーなど非言語情報も大きな役割を果たす。すなわち、人狼プレイヤーの構築の研究は、人工知能を中心に、自然言語、対話、音声、心理学、エージェントなど多様な関連分野の研究とその統合が必要となる挑戦的な課題といえる。実際、人狼知能プロジェクトのオーガナイザーはそうした異なる分野から様々なアプローチで研究を進めている。

人狼知能大会

人狼知能プロジェクトでは、提供している人狼知能プロトコルと人狼知能サーバを用いて、人狼知能大会を開催してきた。

人狼知能大会では参加者を募り、参加者の作成した人狼知能エージェント間で自動対戦を行う。エージェント間では規定された簡易言語であるプロトコルのみを用いて通信する。大会は予選と決勝がおこなわれ、2016年の大会決勝では15エージェントが出場、ランダムに役職を割り振られ、合計で1,124,890ゲームを実行した。プロトコルによる大会部門は今後もプロトコルを発展させつつ継続する予定である。

4 人狼知能大会の自然言語処理部門

2017年の大会より、従来のプロトコル部門に加え自然言語処理部門を設置する。枠組みとしては従来の人狼知能サーバを流用し、投票などゲーム特有のアクションについては従来通り、プロトコルを用いていた「会話」に相当する部分はすべて自然言語のみとする予定である。また当初はプレイヤー数が少なく、結果ターン数も少なくなる設定を考えている。

当面の現実的な目標としては、会話を通じたゲームが成立したといえるレベルを達成するこ

とであろう。既存の対話システムそのままではシステム間の対話がほとんど成立しないであろうから、状況の限定が必要ではないか。実際の人狼ゲームでは雑談的な会話も頻繁に発生し、そこからある種の推理を行うこともあるが、自然言語処理部門ではまずはゲーム勝敗に直結する会話を中心となり、自然と使用する語彙も限定されある程度意思疎通が可能ではないかと期待している。エージェント内でプロトコルを中間言語的に利用することも考えられる。

一般に対話システムの評価は「対話が成立しているように見える」という表層的なものになりがちで、結果ある種の逃げに走るほうが良い評価を得られることも多い。人狼知能においては、会話に制約がないと同時にゲームを成立させ勝利に導くという目標があるため、表層的な発話に終始しては内容の一貫したふるまいにならず不自然に映るだろう。そのため適切な項目を設けた主観評価によって、より本質的な要素に高得点をつける評価が可能ではないだろうか。

対戦形式として、エージェントのみでの対戦と、人間を混在させた対戦があり、いずれも主観評価が妥当と思われる。後者であれば対戦相手として評価することも可能である。対戦という点では勝敗率も一つの評価軸になるが、システムの性能が低い場合はゲームになっておらず、勝敗が適切な評価基準にはならないだろう。人狼自体も勝ち負けだけでなく「面白いプレイ」であったかが重要であり、勝敗以外の評価軸をどう設定するかが研究テーマの一つである。

5 おわりに

本稿では人狼ゲームと人狼知能プロジェクト、過去の人狼知能大会について概説し、新たに開催する自然言語処理部門の構想について述べた。今後大会の開催を通じて、対話システムの本質的な要素の発展に寄与できればと願っている。

謝辞

人狼知能プロジェクトオーガナイザーの鳥海不二夫、大澤博隆、片上大輔、篠田孝祐、松原仁、大槻恭士の各氏に感謝いたします。

参考文献

- [1] 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, Vol.30, No.1, pp.65-73 (2015)
- [2] 鳥海不二夫, 片上大輔, 大澤博隆, 稲葉通将, 篠田孝祐, 狩野芳伸. 人狼知能 だます・見破る・説得する人工知能. 森北出版. (2016)
- [3] “人狼知能プロジェクト.” [Online]. <http://aiwolf.org/>.