

プロトコルを用いた意味理解による

人狼知能の自然言語エージェント

Natural Language AIWolf Agent

by Semantic Understanding Using Protocol

箕輪峻¹ 滝波秋穂¹ 小川ちひろ¹ 三原直樹¹ 真木裕子¹ 柴淳¹ 狩野芳伸¹

Takashi Minowa¹, Akiho Takinami¹, Chihiro Ogawa¹, Naoki Mihara¹,

Yuko Maki¹, Atsushi Shiba¹, and Yoshinobu Kano¹

¹ 静岡大学

¹Shizuoka University

Abstract: The purpose of Artificial Intelligence based Werewolf Project is to construct artificial intelligence that can play "Werewolf". "Werewolf" is a game where players try to find hidden werewolf player through conversations. In this game, each player is assigned a role at the start of the game. The roles include villager, seer, lunatic and werewolf. A player does not know other players' roles. In this situation, non-werewolf players are required to find the werewolf player mainly through conversations. Therefore, the conversations between players are very important. AIWolf Project has two departments: the protocol department and the natural language department. Players can utter only some limited protocols in the protocol department, while players are allowed to talk anything in the natural language department. This paper describes our agent system for natural language department that performs semantic understanding by our own protocols. Our system converts an input text to our protocols based on results of syntactic analysis and case-role analysis, infer other players' roles, and generates sentences.

1. はじめに

近年、我々の日常生活において自然言語をインターフェースとして利用するシステムが増加している。これらによって我々は音声あるいはテキストで入力するだけで自動的に目的のアプリケーションを起動し必要な情報を取得することができるようになった。それだけでなく雑談に対応したシステムも存在し、空き時間に暇つぶしの相手として利用することができる。このように自然言語をインターフェースとするシステムは急速に普及しつつあるが、その一方で現状ではこれらには例えば過去の対話との整合性や文脈の理解など、自然なコミュニケーションを行うために解決しなければならない問題が多数残されている。

コミュニケーションゲームである「汝は人狼なりや (以下人狼ゲーム)」について、篠田ら[1]は人工知能の標準問題として人狼ゲームを用いることを提案している。人狼ゲームとはプレイヤーに対してラン

ダムに割り振られる役職のうち「人狼」と呼ばれる役職のプレイヤーをコミュニケーションによって探し出すゲームである。役職には「村人」「占い師」「狂人」「人狼」などが存在し人狼のプレイヤーの誰が人狼なのか互いに知ることができるが、それ以外のプレイヤーは互いの役職を知る事ができない。人狼ゲームは「昼」「夜」の2つのフェイズを繰り返す事によって進行し、人狼ゲームでは上記2フェイズをあわせて「1日」ないしは「n日目」のように表記する。昼フェイズではプレイヤーは「～が人狼だと思う」「～に投票しよう」「おはよう」のように自由に意見の表明や雑談を行うことができる。夜フェイズでは全プレイヤーの投票により1名を選択する処刑、人狼のプレイヤーの投票により1名を選択する襲撃などのアクションが行われ、処刑および襲撃で選択されたプレイヤーは「死亡」となりゲームから除外される。また、占い師という役職のプレイヤーは自分以外の1名を選択し、そのプレイヤーが人狼かどうか知ることができる (これを「占い」と呼ぶ)。これ

を繰り返し、人狼のプレイヤーを全て処刑できれば村人チーム（「人狼」「狂人」以外の役職のプレイヤー）の勝利となり、生存しているプレイヤーの半数以上が人狼のプレイヤーとなれば人狼チーム（「人狼」「狂人」の役職のプレイヤー）の勝利となる。人狼ゲームにおいてプレイヤーは自由に発言を行うことができるが、その一方で人狼を見つけ出すという目標が発話内容に制限を与えている。このような性質から人狼ゲームはタスク指向と雑談対話の中間的な性質を持つとすることができる。また人狼ゲームをプレイするためには過去の発話との整合性を保つことも要求されることから現状の対話システムが抱える問題を解決するためのタスクとして適していると考えられる。

人狼ゲームをプレイする人工知能の開発を目的としたプロジェクトとして人狼知能プロジェクト[2]が存在する。人狼知能プロジェクトには人狼ゲームにおいて頻繁に用いられる発話をモデル化した人狼知能プロトコル[3]を用いて対戦を行うプロトコル部門が存在し、過去3回大会が行われている。また2017年から人狼知能プロトコルではなく自然言語による自由な発話でゲームを行う自然言語部門が開始され、GAT2017にてプレ大会[4]、CEDEC2017[5]において第1回大会が行われた。本論文では自然言語部門第1回大会に参加し技術賞を受賞したエージェントについて述べる。

2. 提案手法

人狼エージェントの機能は主に他エージェントの発話を理解するための発話解析、解析結果を元に役職等を推測する役職推定、自身の持つモデルや状況から発話を生成する発話生成からなる。本論文では上記3種に加えてアバターの感情表現等を制御するアバター制御を含めた4種について述べる。

2.1. 発話解析

発話解析では入力として与えられた自然言語文を解析し、本システムが利用可能なプロトコルへと変換する。人狼ゲームには一般的な発話には存在しない用語が存在するため、解析を行う前にこれらに対して言い換えを行う。また、解析器としてJUMAN・KNP[6]を使用している。KNPの係り受けや格解析は言い回しによって解析結果に差異が生じることがあるため、そのような言い回しの統一も行う。下表1に言い換えルールの一部を示す。下表の左が言い換え前、右が言い換え後である。

表 1. 言い換え規則

一人称	私
二人称	あなた
吊る	処刑する
噛む	襲撃する
裏切り者,狂信者	狂人
占師、	占い師

これらのルールに基づいて言い換えを行った後、KNPによる解析を行い、その結果を元にプロトコルに変換する。人狼知能プロトコルは表現できる内容が非常に限定的であったことから、大澤[7]の提案した人狼言語を元に改良を加えたプロトコルを用いる。本システムで用いるプロトコルは106単語の組み合わせによって表層的な表現の違いによらず文の意味内容を表現することを目的とする。右表2は自然言語文から本システムで利用するプロトコルへの変換結果の一部を抜粋したものである。表のNは自然言語文、Pは対応するプロトコル文である。

表 2. プロトコル例

1N.	Agent[01]が村人 CO したのであなたに投票します
1P.	{ { vote you } because { Agent[01] did confess { is villager } } }
2N.	あなたはなぜ Agent[01]に投票したいのですか
2P.	{ WHY { you want vote Agent[01] why } }
3N.	私は明日吊られるだろう
3P.	{ i be execute tomorrow }

自然言語文から上記プロトコル文への変換にはKNPの解析結果を用いる。KNPでは係り受け関係のほか表層格も取得することができるが、入力文が話し言葉であるため頻繁に省略が発生することや、プロトコル文の構造と日本語の構造が異なることから、それらのみを手がかりとして利用するのでは不十分である。例えば「私はAgent[01]にAgent[02]に投票してほしい」はプロトコル文では「{ i request { Agent[01] vote Agent[02] } }」と表記されるが、これらは表層格の情報のみでは変換することが難しい。そのため本システムではプロトコルへの変換を行う際に係り受け関係の他にKNPに付与されるfeature[8]を利用する。

2.2. 役職推定

役職推定ではエージェントと役職可能性の組み合わせについてN-Bestリストを作成する。役職推定の

手がかりとして過去の発話や投票行為等を利用し、それぞれの手がかりについて人手であらかじめ重みを定め、それらの積を利用する。エージェントと役職のある組み合わせ 1 について、各エージェントが各役職である手がかり i の持つ重み m の積を 1 のスコア S とする。

$$S_i = \prod m_i$$

2.3. 発話生成

発話生成では役職推定で作成した N-Best リストや過去の発話、現在の状況を元に発話を作成する。生成する発話には主に意見表明、質問、質問への回答、雑談の 4 種類が存在し、具体的な発話内容は役職によって異なる。最初 (0 日目) の昼フェイズは何も情報がなく、また 0 日目の夜フェイズでは「占い」以外のアクションが行われない。そのため 0 日目は挨拶専用のターンとなっており、0 日目には役職にかかわらず挨拶を 1 度だけ発話する。1 日目以降は役職によって生成する発話が異なり、役職が村人の際の発話生成が基本となる。

役職が村人の場合には、日付 (ターン数) に関わらず 1/6 の確率で質問を行い、それ以外の確率で意見の表明を行う。意見表明では役職推定で作成した N-Best リストを元に「～が怪しいと思う」などの発話を生成する。また、同一ゲーム中に同じ発話が行われると不自然に見えるため 1 度発話した内容は保持しておく、同一発話は行わないようにする。質問および意見表明で生成された発話がすでに発話済みの場合、他のプレイヤーから質問をされていればそれに対して回答する。またこの際、質問内容が「なぜ～」のように理由に関する質問だった場合、N-Best リストに紐付けられている手がかりの内容を元に理由を生成する。他プレイヤーからの質問で未回答のものが存在しない場合は一定確率で雑談を行う。

役職が占い師の場合、1 日目の最初の発話で CO (自らの役職を他プレイヤーに宣言すること) を行い、以降は毎日直前の夜に行った占いの結果を最初に発話する。CO および占い結果の発話が行われた後は村人と同様の処理を行う。人狼および狂人の場合も占い師の場合と同様に 1 日目の最初の発話で占い師騙り (自身の役職が占い師だと CO すること) を行い、以降は占い師のように直前の夜の占い結果を発話する。人狼、狂人は当然占いを行うことができないため、これらは役職推定で生成した N-Best を元に自身が占い師だった場合に怪しいプレイヤーを占ったことにして発話を生成する。

また、本システムではゲーム開始時に 5 種類のキャラクターから 1 種類を選び、役職を問わず発話への

キャラクター付けを行う。キャラクター付与は 0 日目の挨拶では定型文が変化し、それ以外の発話では一人称や二人称の呼び方、語尾が変化する。これらはあらかじめ各キャラクターが使用する一人称や二人称、語尾をルールベースで定めておき、発話生成によって生成されたキャラクター性が付与されていない発話に対して適用することでキャラクター付与を行う。

2.4. アバター制御

本システムでは前述した 5 種類のキャラクターに対して下図 1 の各キャラクターが用意されている。これらのキャラクターは Live2D Cubism[9]を利用して作成されており、システムから任意のタイミングで表情の変換などを行うことができる。

図 1. アバター一覧



これらのアバターは初期状態の他に「笑顔」「疑い」「驚き」「怒り」「困惑」の 5 種類の表情が存在し、本システムでは発話生成で生成した文中にこれらを制御するパラメータを加えることで、発話開始や終了など区切りとなるタイミングのほかに発話中の任意のタイミングで表情の切り替えを行うことができる。またこれらのモーションの表示にはゲームエンジン Unity[10]を使用している。

3. 今後の課題

本論文では人狼知能プロジェクト自然言語部門における自然言語エージェントシステムについて述べた。本システムは人狼知能大会第 1 回自然言語部門において技術賞を受賞しているが、未だ多くの課題が存在する。本システムの大きな問題として全ての処理をルールベースで行っていることがまず挙げられる。テキストベースの人狼ゲームとしてインターネット上で行われる BBS 人狼が存在するが、これらは掲示板形式のため 1 発話が長くなりがちであり、また様々なネットスラングや記号、顔文字、箇条書きなど多分にノイズを含んでいるため学習データと

して利用が困難である。そのため本システムではルールベースを採用したが、ルールベースで全ての発話に対応することは不可能である。したがって言い換え処理やプロトコルなどについて人手でルールを作成するのではなく自動的にルールを獲得する方法を検討する必要がある。また同時に現在採用しているプロトコルについて、単語数やプロトコルの文法規則が妥当なものかどうかの検証も必要である。同様に発話生成についても、自然な発話を実現するためには定型文ではなくある程度の柔軟性を持たせる必要がある。

また人狼ゲームはルール上、多人数対話が前提となるが、プレイヤーが全て本システムの場合と本システムと他のシステムが混在する場合、人間のプレイヤーが存在する場合で行われる発話は全く異なる。その一方で自然言語部門大会が現在第1回であることやシステムを公開しているチームが少ないことから、複数システムが存在する際の発話について評価を行うことは難しい。したがって評価は本システムのみ、あるいは本システムと人間のゲームのみが主になるが、これらの評価のみでシステム全体の評価を十分に行うことができるとは考えにくい。そのため本タスクにおける評価方法についても検討が必要である。

謝辞

本システムのアバターデザインおよびモーションは宝塚大学東京デザイン学部渡邊哲意研究室にご提案いただいたものである。深謝申し上げる。また、広島市立大学の稲葉通将氏をはじめ、大会運営の人狼知能プロジェクトオーガナイザー各氏に感謝申し上げます。加えて、本研究の一部は科研費若手研究(A)26700019 および挑戦的萌芽研究 26580082 の助成を受けたものである。

参考文献

- [1] 篠田孝祐, 鳥海不二夫, 稲葉通将, 大澤博隆, 片上大輔: 人工知能標準問題としての人狼ゲームの提案, インテリジェントシステム・シンポジウム講演論文集, 2014(24), 74-77. 一般社団法人日本機械学会, (2014)
- [2] 人狼知能プロジェクト Artificial Intelligence based Werewolf : <http://aiwolf.org/>, (2017)
- [3] 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, 森北出版株式会社, (2016)
- [4] 箕輪峻, 狩野芳伸, 柴淳, 黒田和矢: 係り受け関係を利用したルールベースによる自然言語人狼エージェントの構築, 第2回 GAT2017/2nd Game AI Tournament 2017, (2017-03)
- [5] Computer Entertainment Developers Conference2017 (CEDEC2017) : <http://cedec.cesa.or.jp/2017/>, (2017)
- [6] 笹野遼平, 河原大輔, 黒橋貞夫, 奥村学: 構文・述語好構造解析システム KNP の解析の流れと特徴, 言語処理学会 第19回年次大会, pp.110-113, (2013-03)
- [7] 大澤博隆: コミュニケーションゲーム「人狼」に於けるエージェント同士の会話プロトコルのモデル化, HAI シンポジウム 2013, (2013)
- [8] 河原大輔: KNP で付与される feature 一覧, http://nlp.ist.i.kyoto-u.ac.jp/index.php?plugin=attach&refer=KNP&openfile=knnp_feature.pdf, (2017)
- [9] Cubism 3 | Live2D : <http://www.live2d.com/ja/products/cubism3>, (2017)
- [10] Unity - Game Engine : <https://unity3d.com/jp>, (2017)