

# 人間と戦える人狼知能に関する考察

窪田 秀哉 福田 龍樹  
北九州工業高等専門学校生産デザイン工学科

## 1. はじめに

近年、人工知能は将棋やチェスといった完全情報ゲームにおいて人間を凌駕するような能力を発揮している。そこで、人工知能の次のステージとして期待されているのが「人狼ゲーム」というものである。現在、人狼知能プロジェクトより、人狼知能と呼ばれる人狼ゲームをプレイ可能なエージェントのプラットフォームが公開されている。そこでは、人狼知能同士が自然言語ではなく規定のプロトコルによって相互に意思疎通を行なっている。そのため、現状の人狼知能はあくまで人狼知能同士の戦いであり、人間と意思疎通が可能とは言えない。

本研究では、既存の研究を基に人間と戦う人狼知能を作成するためにエージェントに必要な機能や、既存プロトコルの改善案について考察した。

## 2. 既存プロトコル・ルール

人狼知能による人狼ゲームでは、表1に示したプロトコルに沿って意思の疎通が行われる。

表1 会話メソッドで用いるプロトコル [1]

種類	表現	意味
推測	ESTIMATE [target] [role]	[target] を [role] と推測
カミングアウト	COMINGOUT [target] [role]	[target] が [role] と公表
占い行為	DIVINATION [target] [speacies]	[target] が [role] と言う 占い結果を公表
霊媒行為	INQUESTED [target] [speacies]	[target] が [role] と言う 霊媒結果を公表
護衛行為	GUARDED [target]	[target] を護衛したことを公表
投票	VOTE [target]	[target] に投票する
賛成	AGREE [utterance]	[utterance] の意見に賛成
反対	DISAGREE [utterance]	[utterance] の意見に反対
会話の終了	Over	もう話すことがないときに使用
様子見	Skip	今は話すことがないときに使用

各人狼知能は、このプロトコルに基づいた他の人狼知能の発話履歴などを解析することでゲームの状況を認識している。

## 3. プロトコルの拡張

人狼ゲームを行う上での必要なアクションに関しては基本的な部分に関しては既存のプロトコルで需要部分はカバー出来ている。一方で、そのアクションをどのような経緯で至ったのかという理由を表示することが不可能である。

「人狼知能vs人狼知能」で対戦を行う場合、作成するアルゴリズム及びプロトコルの複雑化、新規参入の高難易度化と言った問題があげられるが、人間と戦うことを想定した場合は理由の解釈を省くことにより容易な実装が可能と推測できる。

## 4. 自然言語処理

人間と対戦することを想定した場合は、前述したプロトコルによる情報のやり取りでは人間は人狼知能に自分の意思を伝える事ができなく、発言の意図からゲームの状況を把握することさえ困難である。

その為、人狼知能が自然言語を発するよりも人間の発言を自然言語処理することよりも人狼知能が発言の意味を理解するシステムを開発する事の方が優先度が高いと考えられる。

## 5. 今後の研究

今回の研究で行ったものは全て仮説であり、実際の検証はまったく行なっておらず、推測の域を出ていない。今後の展望としては既存の高勝率な人狼知能と人間と戦えるようにするために実際の人間の発言を自然言語処理によって現在人狼知能に用いられているプロトコルに変換するシステムを開発することで既存の人狼知能でも人間と同等に戦えるような環境を構築しようと考えている。

## 参考文献

- [1] 大澤博隆, 稲葉通将, 篠田考祐, 狩野芳伸, 片上大輔, 鳥海不二夫: 人狼知能: だます・見破る・説得する人工知能, 森北出版, 2016