

人狼ゲームにおける深層強化学習を用いたエージェント

福田 宗理[†] 穴田 一[†]

東京都市大学[‡]

1. はじめに

近年、人工知能による将棋や囲碁などのゲームの大会が開催されている。その1つに人狼ゲームを行う人狼知能大会[1]がある。人狼ゲームで勝つためには嘘をつく能力や、情報の真偽を見極める能力が要求される。これらの能力が向上することで、人工知能がより高度な判断が可能となり、人間の知能に近づくと考えられる。そのため、この人狼知能大会が注目され、人狼知能の研究[2][3][4][5][6]が勢力的に行われている。

そこで本研究では、役職推定者の役職や全会話情報を用いた役職推定モデルを構築し、その有効性を確認した。そして、この役職推定モデルと Deep Q Network (以下, DQN) を用いて相手から自分への発言内容を状態として入力し、勝利に貢献する為の発言や行動を行うエージェントを構築し、性能評価を行った。

2. 人狼大会のルール

人狼ゲームは村陣営と狼陣営に分かれて、勝ちを競うゲームである。村陣営は人狼を全てゲームから除外する(以下, 追放する)こと、狼陣営は、人間の数を人狼の数以下にすることが勝利条件である。

人狼知能大会では5人狼ゲームと15人狼ゲームが行われていて、本研究では15人狼を取り扱う。15人狼の場合、各プレイヤーは役職(村人, 占師, 霊媒師, 狩人, 人狼, 裏切り者), 種族(人間, 狼), 陣営(村, 狼)の属性を持つ。またプレイヤーの役職と種族, 陣営は人狼同士を除き, 他のプレイヤーに非公開である。村陣営は全て種族が人間の村人6人, 占師1人, 霊媒師1人, 狩人1人である。村人は能力を持たない。占師は1日1人のプレイヤーの種族を調べる(以下, 占い)ができる。霊媒師は死んだプレイヤーの種族を調べる(以下, 交霊)ができる。狩人は1日1人のプレイヤーを人狼の襲撃から守ることができる。狼陣営は人狼3匹で種族は狼, 裏切り者1人で種族は人間である。

人狼は人狼同士で会話を行える。また, 全員で1日1人のプレイヤーをゲームから除外する事を試みる(以下, 襲撃)。裏切り者は狼陣営が勝った時に共に勝利する特殊な役職である。

人狼ゲームには, 昼と夜のフェイズが存在する。昼はプレイヤー同士で会話し, 最後に追放投票を行う。

夜は占師と霊媒師, 狩人, 人狼がそれぞれ固有の行動を行う。また人狼ゲームは占師が一度占いをを行い, ゲームを開始する。

プレイヤー同士の会話はターン制である。発言は定型文で1日に10回まで可能で, 発言をしないことも可能である。全員が1回ずつ定型文を発言することを1ラウンドとし, これを複数ラウンド繰り返す。発言する順番はラウンド毎に無作為に決定する。発言に用いる定型文の種類の一例を表1に示す。

表1: 発言の定型文

発言	意味
ESTIMATE P R	PはRだと思う
COMINGOUT P R	PはRであると宣言する
DIVINATION P	Pを占う
GUARD P	Pを守る
VOTE P	Pに投票する
ATTACK P	Pを攻撃する
DIVINED P S	Pの占い結果はSである
IDENTIFIED P S	Pの交霊結果はSである
GUARDED P	Pを守った
VOTED P	Pに投票した
REQUEST P T	PにTをして欲しい
INQUIRE P T	PにTの内容を質問
BECAUSE T1 T2	T1だからT2だと主張
DAY X T	X日にTだった
NOT T	Tでは無い
AND T T...	複数のTである
OR T T...	複数のTのうち, 少なくとも1つは正しい
XOR T1 T2	T1かT2である
SKIP	パス
OVER	今日はもう喋らない

Pはプレイヤー, Rは役職, Sは種族, Tは定型文, Xは数字を表している。

3. 提案手法

本研究では, 15人狼で勝つために, 役職推定モデルとそのモデルを搭載したエージェントを提案する。

3.1 役職推定モデル

役職推定モデルでは, NNを用いて, 経過日数と役職推定者の役職, 発言情報のうち, 全てのCOMINGOUTの発言と他の新しい発言の計100個, 各プレイヤーの生死をプレイヤーの特徴として入力し, 推定対象である

Artificial Intelligence for Deducing Roles of Players in the Werewolf Game using Information about Conversations among Players

[†]Munemichi Fukuda

[‡]Graduate School of Integrative Science and Engineering, Tokyo City University

[†]Hajime Anada

プレイヤーの役職を推定するモデルを提案する。発言情報は1つの発言について、日にちとターン数、誰が、誰に、発言の種類、役職(一部の発言のみ)、種族(一部の発言のみ)の情報を記録している。

3.2 人狼エージェント

本研究で提案した役職推定モデルとDQNを用いて勝利に貢献する為の発言や行動を行うエージェントを構築する。

本研究のDQNでは報酬としてゲームに勝利した場合に+1、敗北した場合は-1を与える。状態として他のエージェントから自分に対する発言とその発言者の特徴を入力する。

また行動としてゲームのフェイズに合わせて発言の方向性または追放投票の対象選択や夜の行動(占い、護衛、襲撃の対象選択)を決める。ここで発言の方向性とは、同じ目的の為の様々な発言を集約したものである。最終的な発言内容は方向性に沿ってルールベースで決める。

このように2段階で発言を決定する理由は、発言の種類を保ったまま、状況と行動の組み合わせを減らすためである。DQNの欠点の一つに状況と行動の組み合わせが多いと学習が進まない点がある。そのため組み合わせを減らす為に、発言の種類を減らす必要がある。しかし単純に発言の種類を制限すると発言がワンパターンになり、発言からどの役職のエージェントかを推定することが容易になる。その為同じ目的の為の様々な発言を集約したものを発言の方向性と定義する事で行動の種類を減らし、発言の方向性に従って最終的な発言内容をルールベースで決めることで様々な発言が可能となる。

入力内容(状態)は下記の4種類である。

- 日にち
- ターン数
- 直前に自分に言われた発言内容
- その発言者の特徴

発言の特徴の詳細を下記の表2に示す。

表2: 発言者の特徴の詳細

推定モデルで推定した役職
自分と意見が合う数
発言者の発言数

DQNの行動として出力するゲーム内での行動の一例を下記に示す。

- 最多票に合わせて投票
- 役職推定モデルの結果で襲撃
- 占師COしているプレイヤーを占う

またDQNの行動として出力するゲーム内での行動の一例を下記に示す。

- 人狼を排除する発言
- 無口のプレイヤーを排除する発言

さらに下記の表3に具体的な発言の一例を載せる。

表3: 具体的な発言の一例

発言の方向性	具体的な発言
人狼を排除する発言	人狼だと推定する
	人狼への投票発言

無口を排除する発言	人狼への投票
	占いを希望する
	無口への投票発言
	無口への投票
村人を信じる発言	無口に誰が人狼だと思うかを質問
	役職推定モデルを信じる
	村の多数意見に賛成
	占い師の占い結果を信じる

4. 結果

役職推定モデルを搭載した人狼エージェントを過去の人狼大会に出場したエージェント5種類とサンプルエージェント1種類9体と2000回対戦して学習を行い、500試合してテストした結果、勝率は39.6%、生存率は37.0%であった。また役職別の勝率は以下の表7の通りである。

表7: 役職毎の勝率

役職	勝率	勝利数/役職回数
村人	27.3%	72/264
占い師	25.8%	8/31
霊媒師	34.2%	13/38
狩人	27.6%	8/29
人狼	73.3%	74/101
狂人	62.2%	23/27

村陣営の勝率が低く人狼陣営の勝率が高くなる。理由は、サンプルエージェントが役職を判別する能力が低いため、村陣営になった時に人狼を探せないが、人狼陣営になった時には仲間の人狼を知っているため、村人を襲撃する事が容易だからである。

5. 今後の課題

現在のモデルでは相手から質問を受けた際に適切に返事を返すことが出来ない為、学習によって返事を返すような報酬の与え方を検討したい。

参考文献

- [1] 人狼知能プロジェクト <http://aiwolf.org/>
- [2] 大川貴聖, 吉仲亮, 篠原歩: “深層学習を用いて役職推定を行う人狼エージェントの開発” 第22回ゲームプログラミングワークショップ(2017)
- [3] 萩原誠, 伊藤孝行, アーメッドムスタファ: “Q学習と役職推定に基づく人狼知能エージェントの作成” 情報処理学会論文誌 vol.60 No.10 (2019)
- [4] 梶原健吾, 鳥海不二夫, 稲葉通将: “人狼知能大会における統計分析とSVMを用いた人狼推定を行うエージェントの設計” 第30回人工知能学会(2016)
- [5] 家原瞭, 廣田敦士, 田中一品, 荒木雅弘, 岡夏樹: “Differentiable Neural Computerを用いた人狼知能の開発” 第32回人工知能学会(2018)