

# 15 人狼ゲームにおける会話情報による役職推定

福田 宗理† 穴田 一†

東京都市大学大学院†

## 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]が勢力的に行われている。その中にニューラルネットワーク(以下、NN)を用いて、役職推定に成功している大川らの研究がある。しかしそのモデルには人狼の推定精度を向上できる余地があると考えられる。何故なら、NNに入力する特徴として、役職推定者の役職や占いの対象となっているプレイヤーとそのプレイヤーの発言など、用いていない情報があるからである。そこで本研究では、役職推定者の役職や全会話情報を考慮し、役職推定を行う。

## 2. 既存研究

### 2.1 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、処刑によって人狼を全てゲームから追放することが勝利条件であり、狼陣営は、襲撃によって人間をゲームから追放して人間の数を人狼の数以下にすることが勝利条件である。

人狼知能大会では人狼ゲームを5人か15人で行い、本研究では15人狼を扱う。15人狼では、各プレイヤーは陣営(村, 狼), 種族(人間, 狼), 役職(村人, 占師, 霊媒師, 狩人, 人狼, 裏切り者)の属性を持つ。またプレイヤーの陣営と種族, 役職は人狼同士を除いて他のプレイヤーに非公開である。

15人のうち、村人8人, 占師1人, 霊媒師1人, 狩人1人は村陣営に属する。残りの人狼3匹と裏切り者1人は狼陣営に属する。役職が村人, 占師, 霊媒師, 狩人, 裏切り者であるプレイヤーの種族は人間である。

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。霊媒師は処刑された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。人狼は全員で1日1人のプレイヤーを襲撃できる。また襲撃前に人狼同士で会話ができる。裏切り者は能力を持たない。

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。狼陣営は時折占師を騙るなどの嘘の発言をしてプレイヤーを騙し、都合の良い展開に持ち込もうとする。村陣営は狼陣営の嘘を見抜き、逆に人狼を見つけようとする。

追放投票は全員で投票を行い、最多票のプレイヤーが追放される。夜は占師が1人を占い、霊媒師が1人と交霊し、狩人が1人を守り、人狼が全員で1人を襲撃する。また、人狼ゲームはゲーム開始直後に占師が一度占いを行い、その後昼から始まる。

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回まで可能であり、発言をしないこと(SKIP, OVER)も可能である。全員が1回ずつ定型文を発言するかSKIP, OVERすることを1ラウンドとし、これを繰り返す。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
ATTACKED P	Pを襲撃した
AGREE T	Tに賛成する
DISAGREE T	Tに反対する
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

Artificial Intelligence for Deducing Roles of Players in the Werewolf Game using Information about Conversations among Players, †Munemichi Fukuda, †Hajime Anada, †Graduate School of Integrative Science and Engineering, Tokyo City University

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.

以下の 3 つの内どれかの条件を満たすと会話が終了し, 追放投票に移行する.

- 全プレイヤーが OVER をする
- 全プレイヤーが SKIP をするラウンドが 3 回連続する
- 20 ラウンドが経過する

## 2.2 大川らの役職推定モデルと人狼推定モデル

人狼ゲームに強い人工知能を作る為には, プレイヤーの会話や占い結果の情報から誰が人狼かを精度よく推定する事が重要である. 大川らの研究では, NN を用いて以下の表 2 に示す 1 人のプレイヤーの特徴から 1 人のプレイヤーの役職を推定する役職推定モデルと, 以下の表 2 に示す 1 人のプレイヤーの特徴を全員分用いて人狼であるプレイヤーを推定する人狼推定モデルを構築した. 役職推定モデルはプレイヤーの役職が村人, 占師, 人狼, 裏切り者である確率をそれぞれ出力する. 人狼推定モデルは各プレイヤーの人狼である確率を出力する.

表 2: プレイヤー X の特徴

特徴	詳細
経過日数	現在何日目か
占師の数	占師 COMINGOUT (以下 CO) をしたプレイヤー数
被占い結果	X が人間判定された数と人狼判定された数
何番目の占師	X が何番目に占師 CO をしたか (占師 CO をしたプレイヤーのみ)
占い結果	X が報告した人間判定の数と人狼判定の数 (占師 CO をしたプレイヤーのみ)
投票変更数	X が VOTE の対象にしたプレイヤーと, X が行った投票の対象が異なった回数
生死	X の状態が生存, 追放された, 襲撃されたかのいずれか
肯定的意見の数	X が別のプレイヤーに対して村陣営であると推定した, または別のプレイヤーの発言を AGREE した数
否定的意見の数	X が別のプレイヤーに対して狼陣営であると推定した, または別のプレイヤーの発言を DISAGREE した数

## 2.3 実際の行動

推定モデルの結果を受けて, 自分が村陣営の場合, 役職推定モデルで人狼だと推定したプレイヤーが 1 人で, かっ生きていればそのプレイヤーに投票する. しかし, 複数いる, または 1 人もいなければ人狼推定モデルで 1 番人狼の確率が高いプレイヤーに投票する. それにより決定した対象が死んでいて投票出来ない場合は無作為に投票する.

また自分が狼陣営の場合, 投票時と襲撃対象の選択時に役職推定した結果, 占師がいればそのプレイヤーを選択する. しかし, 推定したプレイヤーに生存している占師がいなければ村人を選択する. しかし, 推定したプレイヤーに生存している村人もいなければ裏切り者を選択する. それにより決定した対象が死んでいて選択出来ない場合は無作為に選択する.

## 2.4 既存手法の問題点

既存手法では, NN に入力するプレイヤーの特徴の中に, 役職推定者の役職や占い対象者となっているプレイヤーとそのプレイヤーの発言など, 用いていない情報があることが問題点である. このことが, 推定精度が良くない主要な要因であると考えられる.

## 3. 提案手法

本研究では, 15 人狼で勝つために, 経過日数と自分の役職推定者の役職, 各プレイヤーの生死, 発行情報を COMINGOUT の発言と他の新しい発言の計 100 個をプレイヤーの特徴として入力した役職推定モデルを提案する. 各ゲームにおいてランダムに 1 人のプレイヤーを役職推定者として選択し, そのプレイヤーの視点でこのモデルを用いて別のランダムに選択した 1 人のプレイヤーの役職を推定する.

発行情報は 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(発言の種類によっては種族, 役職)の情報を記録している.

人狼知能大会のログデータを学習データに用いる. 提案手法をゲーム中の区切りごとに用いて日にちごとの正答率の変化を調べて考察する. パラメータは以下の表 3 の通りである.

表 3. パラメータ

	提案手法
入力次元	7765
出力次元	6
中間層の数	2
中間層のノード数	500
活性化関数	ReLU

提案モデルによる出力結果とテストデータの正解ラベルを比較して, 精度と再現率を確認する. 精度は, 推定モデルが推定した結果の内, 正解ラベルが一致している割合である. 再現率は, 正解ラベルの内, 推定モデルによる推定が一致している割合である.

実験結果の詳細と考察は発表時に述べる

## 参考文献

- [1] 人狼知能プロジェクト <http://aiwolf.org/>.
- [2] 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” The 30th Annual Conference of the Japanese Society for Artificial Intelligence, (2016).
- [3] 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” The 22nd Game Programming Workshop (2017).
- [4] 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).