

阪本 真基 下山 勝之 加守田 侑 上野 敦志 田窪 朋仁  
 大阪市立大学 大学院 工学研究科 / 工学部

## 1. はじめに

我々は図1に示すように人狼BBS[1]のログを取得して分析することで役職を推定する基準を作成し、その基準をもとに人狼ゲームを行うエージェントを作成した。以下では、それぞれの過程を述べていく。

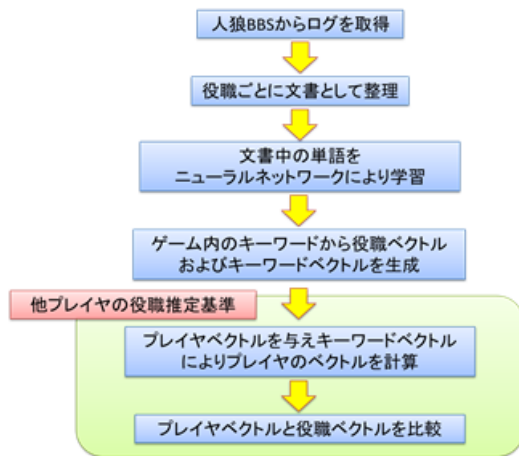


図1: 役職推定基準作成の流れ

## 2. 人狼BBSによるログの取得

我々は人狼を行うエージェントを作成するにあたり、ゲーム内の意思決定を定めるために、インターネット上で人間同士が実際に人狼ゲームを行っている人狼BBSのログの分析を行った。

まず、分析を行う上で役職ごとにどのような発言が行われているか把握するため、役職ごとに整理を行った。人狼BBSで行われるゲームにはプロローグ、1日目、2日目、…、エピローグがある。エピローグでは各プレイヤーがどの役職であったかが公開される。そこで、エピローグからプレイヤーと役職の対応表を獲得し、役職ごとの発言内容を文書としてまとめた。

## 3. ログの分析・ベクトルの生成

ログ内の発言は人間が話す自然言語であるため日本語においては分かち書きが必要である。そこで、各文書においてmecab[2]による分かち書きを行った。しかし、人狼BBSでよく用いられる「霊能者」や「非占霊」は「霊、能、者」「非、占、霊」と分解しすぎてしまうので予め辞書登録をする。

分かち書きを行った文書から単語を抽出する。ここで、単語とは「名詞」かつ「漢字または大文字半角英字」を含むものとした。これは人狼BBSにおいてひらがなやカタカナのみを用いた単語が少なく重要度が低いこと、英字には「GJ」や「CO」といった単語はほしいが「www」のような単語はいらぬためである。

文書ごとに抽出された単語を用いてニューラルネットワークによる学習[3]を行う。これにより、単語の意味の類似性を考慮したベクトル表現を獲得できる。例えば、「CO」に近い単語として「対抗」や「潜伏」が、「GJ」に近い単語として「狩人」や「襲撃」が得られる。

## 4. キーワードベクトル・役職ベクトルの生成

人狼知能大会で用いられる発言は「ESTIMATE」「COMINGOUT」「DIVINED」「INQUESTED」「GUARDED」「VOTE」「AGREE」「DISAGREE」「ATTACK」「SKIP」「OVER」である。これらと学習によって得られた単語ベクトルから対応したベクトルを得る。しかし、「SKIP」「OVER」は発言なし、「ATTACK」は人狼会話でしか使用されないため無視する。「AGREE」「DISAGREE」「ESTIMATE」については対応するベクトルは得ず、別の方法を用いる。表1にそれぞれの発言に対応する単語を示す。複数挙げられているものはその単語ベクトルの平均値を用いる。これによって得られたベクトルをキーワードベクトルと呼ぶ。

役職ベクトルは各役職において、表1で示した単語を使用した回数からベクトルの平均値を求め、それを用いる。例えば、占い師が「CO」を1000回、「占」を5000回使用していた場合、次式で定義される。

$$\text{占い師ベクトル} = \frac{1000 \times \text{「CO」ベクトル} + 5000 \times \text{「占」ベクトル}}{6000} \quad (1)$$

得られたキーワードベクトル及び役職ベクトルを図2、図2の原点を拡大したものを図3に示す。この図2および図3は「Villager」を原点に取っており、役職ベクトルである「Villager」「Seer」「Medium」「Bodyguard」「Possessed」「Werewolf」は10倍の大きさのベクトルで表している。

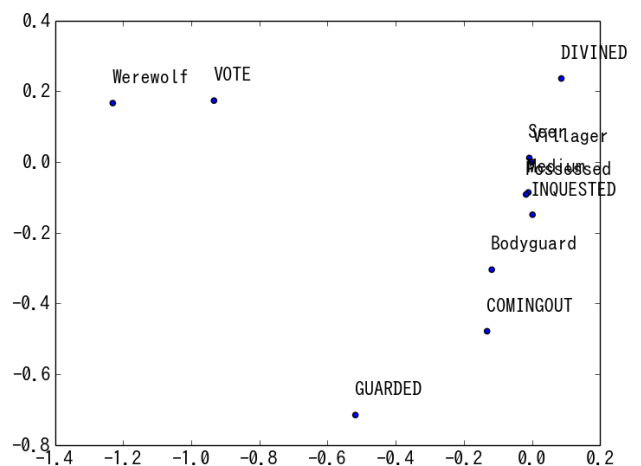


図2: キーワードベクトルおよび役職ベクトル

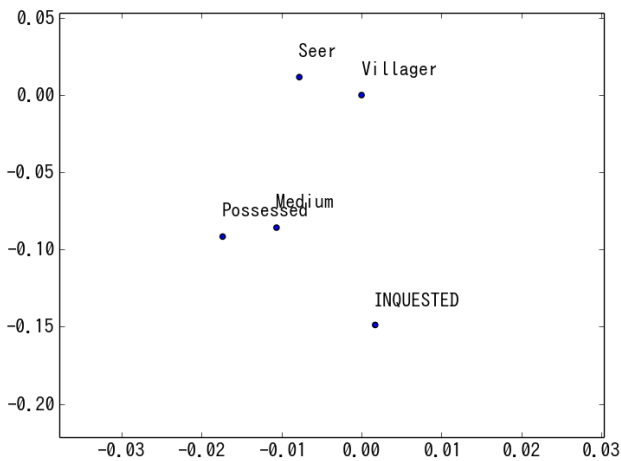


図 3: キーワードベクトルおよび役職ベクトル (拡大)

図 2 および図 3 を見ればわかるように「Werewolf」には「VOTE」が、「Seer」には「DIVINED」が、「Bodyguard」には「GUARDED」が、「Medium」には「INQUESTED」が近くにある。

表 1: 発話に対応する単語

キーワード	単語
COMINGOUT	CO
DIVINED	占い、占
INQUESTED	霊、霊能
GUARDED	GJ
VOTE	投票、処刑

## 5. ベクトルによる意思決定

キーワードベクトルおよび役職ベクトルが得られたら、各プレイヤーに現在どの点にいるかというベクトル (以下、プレイヤーベクトル) を与える。プレイヤーベクトルは役職ベクトルの平均値で初期化する。

図 4 に示すようにプレイヤーが発話を行う度に、そのキーワードベクトルを取得し、キーワードベクトルへ向かうように発話者のプレイヤーベクトルに加算する。プレイヤーベクトルを  $P$ 、キーワードベクトルを  $W$  とすると次のようになる。

$$P = P + \alpha (W - P) \quad (2)$$

ここで、 $\alpha$  は制御パラメータであり、どれだけそのキーワードベクトルに向かっていくかである。ここでは、1 日の最大発話数が 10 であることから  $\alpha=0.1$  としている。

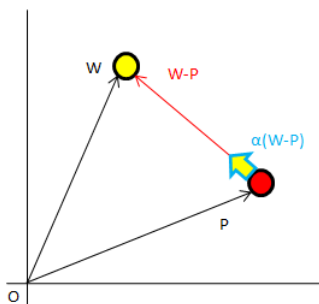


図 4: プレイヤベクトルの移動

「AGREE」「DISAGREE」はその対象発話の発話者を取得し、発話者のプレイヤーベクトルを  $O$  とすると、「AGREE」

の場合は  $O$  に近づくように、「DISAGREE」の場合離れるようにプレイヤーベクトル  $P$  を更新する。また、「ESTIMATE」はその目的語が「Werewolf」の場合、「VOTE」と発話した場合はその対象のプレイヤーのベクトル  $O$  から離れるように更新する。「DISAGREE」「ESTIMATE」「VOTE」の場合、プレイヤーベクトルの移動は次のようになる。

$$P = P - \alpha (O - P) \quad (3)$$

相手プレイヤーの発話を受けて更新していくことで、プレイヤーベクトルはどこかの役職へ近づいていく。これにより、プレイヤーベクトルと役職ベクトルの比較を行うことで、役職の推定基準とする。例えば、人狼と思われるプレイヤーは Werewolf ベクトルに最も近いプレイヤーとする。

## 6. 各役職の戦略

今回の人狼知能大会では、この役職推定基準を以下のように用いた。村人、霊能者、狂人は能力の使用に役職推定を用いていない。

- ・狩人「Seer」に最も近いプレイヤーを守る。
- ・占い師「Werewolf」に近いプレイヤーを占う。
- ・人狼「Bodyguard」「Seer」の優先順位で襲撃する。

共通の発話処理としては、不都合な発話を受けた場合の否定と人狼推定を行う。しかし、人狼側の役職は「Seer」に近いプレイヤーを人狼として推定する。

投票は発話により「VOTE」の対象が 1 人に集中している場合と自分の人狼推定を使い分けるために、その日の各プレイヤーが「VOTE」と言われた回数を記憶しておき、最も多く「VOTE」と言われたプレイヤーの回数を  $vote\_max$ 、「VOTE」の発話回数の合計を  $vote\_sum$  とすると、 $voted\_max/voted\_sum$  の確率で最も多く「VOTE」と言われたプレイヤーに投票する。しかし、 $vote\_sum$  が 10 以上の場合に限る。そうでない場合は、ベクトルにより人狼推定を行い、最も人狼に近いプレイヤーに投票を行う。

占い師は発話により「VOTE」と言われた回数が最も多いプレイヤーが「Werewolf」に最も近いプレイヤーの場合、2 番目に近いプレイヤーを占う。

## 7. おわりに

今回の人狼知能大会では、役職を推定する基準を作成するために、事前学習として人狼 BBS のログを取得し分析を行った。そして、推定基準をもとに各役職の能力の対象の決定や人狼の推定を行うエージェントを作成した。

今後の課題としては、エージェントとの対戦だけでなく、人間を相手した場合の勝率や役職の推定の性能評価を行うことが求められる。

## 参考文献

- [1]<http://www.wolfg.x0.com/>
- [2]<http://taku910.github.io/mecab/>
- [3]Joo-Kyung Kim, Marie-Catherine de Marneffe: "Deriving adjectival scales from continuous space word representations", Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pages 1625-1630, Seattle, Washington, USA, 18-21 October 2013