

人狼知能エージェントの簡易生成システムの開発

Development of a simple generation system of the werewolf AI agent

武田 惇史^{*1} 鳥海 不二夫^{*1}

Atsushi Takeda

Fujio Toriumi

^{*1}東京大学

The University of Tokyo

In recent years, the game "werewolf" has attracted attention as a standard problem replacing shogi and chess in the field of artificial intelligence. The game "werewolf" is a communication game that belongs to the imperfect information game, and the AI that plays werewolf is called "AIWolf". A platform for the AIWolf agents to play the game werewolf has been already created, and "AIWolf Competition" has been held five times to recruit general developers and compete for the strength of the agents they developed. The purpose of the competition is to promote development of AIWolf by collective intelligence. Therefore it is considered important to have more developers participate. A problem that arise is that the number of developers who can participate is limited because the developers are required to be able to write some computer programs. In this paper, to solve this problem, we develop a simple AIWolf generation system that does not require programming by parameterizing AIWolf based on the winning agent of the past competition.

1. 序論

人工知能技術の性能指標として古くから用いられているものとして、ゲームが挙げられる。ゲームは、ルールや勝ち負けが明確であること、一定の複雑さを有すること、人間のプレイヤーとの比較ができることなどの性質を持ち、これら点で人工知能のベンチマークに適しているとされる。人工知能が人間を超えたゲームの例として挙げられるのは、チェッカー [J. Schaeffer 97], チェス [M. Campbell 02], 囲碁 [D. Silver 16] などである。ここで挙げたゲームは、プレイヤーにすべての情報が開示される「完全情報ゲーム」という分類に属するゲームである。一方、プレイヤーに開示される情報が限定的であるようなゲームは「不完全情報ゲーム」と呼ばれるが、これに関する研究はあまり行われていない。

これを受けて、近年では不完全情報ゲームの一種である「人狼ゲーム」が人工知能研究において将棋やチェスに代わる標準問題として注目されており [片上 15], 人狼ゲームをプレイする AI は「人狼知能」と呼ばれる。人狼知能エージェント同士が人狼ゲームをプレイするプラットフォームが作られ、一般の開発者を募ってエージェントの強さを競う「人狼知能大会」が過去に 5 回行われている。

人狼知能研究の最終目的は、人狼ゲームを本質的に理解し、人狼ゲームを自然にプレイできる人狼知能を作成することであり、究極的には高度なコミュニケーション能力をもった対話エージェントの作成につながるものであると考えている。

その目標の第一歩として、集合知による人狼知能開発の促進を目標としているのが人狼知能大会である [鳥海 17]。そのため、大会ではより多くの開発者に参加してもらうことが重要になると考えられる。そこで問題となり得るのが、開発においてプログラミングが必須となっているために参加可能な開発者が限られているという点である。

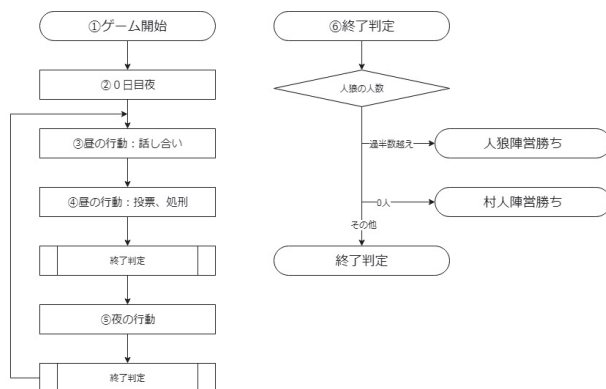


図 1: 一般的な人狼ゲームの進行

本研究では、この問題を解決するため、2019 年に行われた第一回国際人狼知能大会の優勝エージェントに基づく人狼知能のパラメータ化を行うことで、プログラミングを必要としない人狼知能の簡易生成システムを開発する。

2. 人狼ゲーム

2.1 ルール

人狼ゲームは、コミュニケーションによってゲームが進行する不完全情報ゲームであり、プレイヤーは 3 人以上の複数人（通常 5 人から 15 人くらいで遊ばれている）で、「人狼陣営」と「村人陣営」に分かれて戦うチーム戦である。

一般的なゲームの進行を図 2.1 に示す。プレイヤーには最初、「役職」がランダムに割り当てられ、役職ごとに人狼陣営であるか村人陣営であるかが定まる。ゲームは「1 日」という単位を繰り返すことで進行する。1 日は「昼」と「夜」のフェーズがある。

昼のフェーズには、生存プレイヤーによる話し合いが行われたのち、多数決に基づく投票によって一人がゲームから除外される（「追放」と呼ばれる）。

夜のフェーズには人狼（人狼陣営の役職）による「襲撃」が

連絡先: 武田惇史, 東京大学大学院工学系研究科システム創成学専攻, 113-8656, 東京都文京区本郷 7-3-1 工学部 8 号館 526, TEL: 03-5841-6991, E-mail: takeda@crimson.q.t.utokyo.ac.jp

表 1: 5 人狼で使用する役職一覧

役職名	陣営	能力
村人	村人	なし
占い師	村人	指定した一人を占い、その役職が人狼かどうかを知ることができる
霊媒師	村人	前の日に追放されたプレイヤーの役職が人狼かどうかを知ることができる
狩人	村人	指定した一人を人狼の襲撃から守ることができる
人狼	人狼	指定した一人を襲撃し、ゲームから除外できる
裏切り者	人狼	なし

行われる。襲撃では、人狼は任意に選んだ一人をゲームから除外することができる。これを繰り返し、敵陣営を殲滅させた陣営の勝利となる。

2.2 役職

本論文では人狼知能大会のレギュレーション^{*1}に従ってシステムの設計を行う。このレギュレーションにおいて使用される役職の一覧を表 1 に示す。

2.3 戦略

通常、村人陣営は人狼陣営より人数が多く、多数決によって行われる追放において有利である。一方、人狼陣営には襲撃という強力な能力を持っている。また、人狼（役職）は他の人狼がだれかを知ることができる。人狼陣営は少数派のため、投票されないように村人陣営のふりをするのが一般的な戦略である。逆に、村人陣営は自分が村人陣営であることを他の村人陣営のプレイヤーに説得したり、人狼陣営のウソを見抜き、人狼を追放したりすることが目的となる。

一方で、ゲームの進行によっては人狼陣営が生存プレイヤーの過半数になることが起きうる。このようなゲームの状態は PP と呼ばれ、プレイヤーの戦略が大きく変わりうる。例えば、人狼陣営で票を合わせることで、投票の過半数を獲得することができるため、人狼が自分の正体を隠す必要はなくなる。

このように、人狼ゲームには相手の役職やゲームの状態を推測する、相手を騙す、説得するなどといった高度なコミュニケーションを要素として持つという特徴がある。

3. 役職推定

3.1 概要

本章では、簡易生成システムで用いられる役職推定手法について説明する。この手法は、第一回国際人狼知能大会の優勝エージェントが使用したものに基づいている。

3.2 変数定義

説明のためいくつかの変数を定義する。

- R_i は i 番目のエージェントの役職である。 R は確率変数であり、この分布を推定することが役職推定の目的となる。
- X は全プレイヤーへの役職の割り当ての組み合わせを表す。すなわち、 $X = \{R_1, R_2, \dots, R_N\}$ である (N はプレイヤーの人数)。
- i 番目に観測した情報 A_i は、ゲームの状態 C_i 、行動主体 S_i 、行動の種類 V_i 、行動対象 O_i の 4 要素からなるもの

とする。プロトコル上でこれより複雑な情報であっても、何らかの方法でこの 4 要素に分解して扱う。なお、対象エージェントが存在しないような種類の行動であっても、便宜上 O_i にそれを表す値が設定されるものとする。

- I_t は、ある時点 t までにエージェントに公開されているすべての情報を表す。すなわち、 $I_t = \{A_1, A_2, \dots, A_t\}$ である。時刻 t を意識する必要がない場面においては単に I と書く。

3.3 役職推定手法

ベイズの定理により、

$$P(X|I) \propto P(X)(I|X) \quad (1)$$

が成り立つ。ここで、 $P(X)$ は一様分布（ゲーム開始時の役職割り当ては一様ランダム）なので、

$$P(X|I) \propto P(I|X) \quad (2)$$

すなわち、ある特定の役職の割り当てに対して、その割り当てが実際の割り当てに一致する確率は、その割り当てを仮定したうえで観測された情報が実際に得られる確率（尤度）に比例する。

役職推定により求めたいのは、各プレイヤーの役職の確率分布、すなわち $P(R_i|I)$ であった。これは、以下の式によって求められる。

$$P(R_i|I) = \sum_{X \ni R_i} P(X|I) \quad (3)$$

以上から、あらゆる役職割り当てに対する得られた情報の尤度が計算できれば求めたい確率分布が求まる。

尤度を求めるにあたり以下の式を仮定する。

$$P(I|X) = \prod_i P(I|R_i) \prod_{i,j} P(I|R_i, R_j) \quad (4)$$

すなわち、3 人以上の関係を考慮に入れない。

求めたい尤度は以下のように行動の尤度の積として表される。

$$P(I|X) = \prod_{(C_i, S_i, V_i, O_i) \in I} P(V_i|R_{S_i}, R_{O_i}, C_i) \quad (5)$$

ここでは、各行動は独立であることが仮定されている。式から、ゲームの状態と役職の割り当てを仮定した時のプレイヤーが起こした行動の尤度を求めればよい。

これらの行動の尤度を求めるにあたって、実際に観測された頻度に基づいてエージェントの行動の予測を行うものとする。頻度変数 $F_{i,R_i,R_j,V_i,C}$ を、 i 番目のエージェントの役職が R_i

*1 <http://aiwolf.org/control-panel/wp-content/uploads/2019/03/人狼知能世界大会2019レギュレーションVer1.3.pdf>

表 2: 行動一覧

行動	対象・条件	説明
投票	対象エージェントを指定	投票先を決める。
襲撃	対象エージェントを指定, 人狼のみ	襲撃先を決める。
占い	対象エージェントを指定, 占い師のみ	占い先を決める。
霊媒	対象エージェントを指定, 霊媒師のみ	霊媒先を決める。
護衛	対象エージェントを指定, 狩人のみ	護衛先を決める。
発言: CO	対象役職を指定	役職表明を行う。「私は〇〇(役職名)です」等。 嘘の可能性もある。
発言: 投票宣言	対象エージェントを指定	発言で投票先を表明する。 「私は〇〇(エージェント名)に投票します」等。
発言: 推定	対象エージェント及び役職を指定	役職推定の発言をする。 「私は〇〇(エージェント名)がXX(役職名)だ と思う。」等。
発言: 占い, 霊媒結果報告	対象エージェントと結果を指定	占いや霊媒の結果を報告する。 「〇〇(エージェント名)占って黒だった」等。

表 3: 観測から得られるゲームの情報一覧

項目	説明・備考
役職推定結果	第3章で述べた役職推定器の結果を利用する。 あるエージェントがある役職である確率が値として利用可能である。
CO 役職	各エージェントの役職表明状況
勝率	各エージェントのこれまでの勝率。大会標準ルールでは、100 戦の間メンバーが変わらず 記憶が保持されるため、それまでの勝率を計算し保持することが可能である。
予想得票数	過去の投票先や、投票宣言、人狼推定発言をもとに書くエージェントの投票先を予測し、 それに基づいて算出した得票数の予測値
ゲーム内時刻	ゲーム内の日にちと発言ターン数
PP 判定状態	生存プレイヤーの過半数が人狼陣営である状態を PP と呼ぶ。ゲームがこの状態にあるかを判定した結果。

で、 j 番目のエージェントの役職が R_j で、ゲームの状態が C であるとき、 i 番目のエージェントが j 番目のエージェントに対して行動 V を行った頻度と定義する。

このとき、行動の尤度は以下のように計算される。

$$P(V|R_i, R_j, C) = \frac{F_{i, R_i, R_j, C, V}}{\sum_V F_{i, R_i, R_j, C, V}} \quad (6)$$

$F_{i, R_i, R_j, V, C}$ は行動の頻度を数える変数であり、その値はゲームの進行とともに動的に決まる。添え字に行動主体エージェント i を含んでいるのは、特定のエージェントの行動の傾向を学習することで、パーソナライズされた行動の尤度を算出することを可能にするためである。

$F_{i, R_i, R_j, V, C}$ には、初期値として非零の値を設定することができる。これは、事前知識としてどのような行動が起きやすいかを推定に組み込むことに対応する。

4. 簡易生成エージェントのモデル化

4.1 概要

本章では、人狼知能エージェントをパラメータ化し、簡易生成システムに応用する方法を提案する。本章では、パラメータを調整することで多様なエージェントを作成できるようにするために制約が少なく、簡易生成システムへ応用するためにパラメータを変えることでエージェントの振る舞いがどう変わるかが説明可能であるようなパラメータ化を目指す。

4.2 推定に関するパラメータ

エージェントは、観測された情報を何らかの形で処理することでゲームの状態を推定し、行動決定においてそれを利用するものとする。エージェントが推定する項目は以下の3つが考えられる

- 各エージェントの役職
- 潜在的な得票数
- PP 状態の判定

各エージェントの役職については第3章で説明した役職推定器を用いる。具体的なパラメータとしては、3.3 節で説明した頻度変数 F の初期値の設定がある。

潜在的な得票数は、過去の投票、投票発言、人狼推定発言から各エージェントの投票先を予測し、これに基づいて得票数を推定するというものである。これに関しては、パラメータ化する必要はないと考える。

PP 状態の判定は、PP 発生によって戦略を切り替えるために必要な情報であり、その方法としては「生存プレイヤーのうち人狼陣営の割合の期待値が x 以上であれば PP と判定」などが考えられ、閾値 x がパラメータとなる。

4.3 行動に関するパラメータ

本論文においては、エージェントはある時点で可能な行動に対する評価関数の値が最も大きいものを選択し行動するというモデル化を行う。状態 s おける行動 a の評価値を $Q(s, a)$ と書

表 4: 実験結果 (勝率)

名称 (大会順位) \ 役職	裏切者	占い師	村人	人狼	総合
簡易生成システム (-)	0.328	0.716	0.770	0.344	0.588
fisherman(5)	0.272	0.745	0.750	0.376	0.576
takeda(1)	0.313	0.798	0.683	0.263	0.565
hello_wolf(2)	0.255	0.738	0.711	0.311	0.543
GO1DENO(4)	0.245	0.573	0.681	0.119	0.447

くことにする。ある時点 i でエージェントが観測できるゲーム情報を s_i 、エージェントの取りうる行動全体を A_i と書くと、エージェントの取る行動は

$$\arg \max_{a \in A_i} Q(s_i, a) \quad (7)$$

と表せる。ここで、観測可能なゲームの情報 s_i およびエージェントがとりうる行動 A_i は、説明可能性の観点から意味や目的が明確なものに限定するものとする。考慮するエージェントの行動一覧を表 2 に載せる。また、考慮する観測可能なゲームの情報の一覧を表 3 に載せる。観測可能なゲームの情報には、各エージェントの特徴量に関するものが含まれ、この部分に第 3 章で示した役職推定器が用いられている。

4.4 簡易生成エージェントの設計

簡易生成エージェントを設計するにあたり、前節までに定義したエージェントは抽象的すぎるためそのままでは扱いにくい。そこで、エージェントの行動や観測する情報を、実際のゲームにおいてエージェントが考慮すると考えられるものへと制限をかけることで簡易生成エージェントの設計をする。

まず、「役職表明」のような行動対象を持たない項目は、評価値によるものではなく、ルールベースで決定できた方が扱いやすいと考えられる。また、評価関数の具体的な計算方法を指定できるようにする必要がある。分かりやすさのため、評価関数への入力行動対象となるエージェントに関する特徴量に絞り、評価関数はその特徴量の線形和として表されるものとする。

これらを踏まえ、最終的にパラメータとなる項目は以下の通りである。

- 役職推定における頻度変数の初期値
- PP 発生判定の閾値
- CO する役職とそのタイミングに対する確率分布。
- PP 発生時、CO する役職の確率分布。
- 発言、投票、占いなどの対象を決定するための評価関数の重み。(例えば、非カミングアウト者にうち、 $(1.0 \times \text{人狼確率} + 0.1 \times \text{得票数} - 0.01 \times \text{勝率})$ が最も高い人に投票など)

5. 強さの簡易評価

本章では、作成した簡易生成システムの出力するエージェントの強さを見積もるための簡易的な実験を行う。

実験設定としては、第一回国際人狼知能大会における決勝出場エージェント 4 体と今回作成した簡易生成システムを 5 人狼レギュレーションで 10000 回戦わせ、勝率を比較するも

のとする。パラメータは優勝エージェントのアルゴリズム^{*2}を参考に設定した。

結果は表 4 のようになり、強さの観点から決勝進出エージェントに劣らない性能を出すことが可能であることが分かった。

6. 結論

本論文では、第 3 章で人狼知能簡易生成器で用いられる役職推定手法を説明した。また、第 4 章ではエージェントの推定と行動をパラメータ化したうえで、それをさらに具体化し、人狼知能簡易生成器のモデル化を行った。

今後の研究として、実際に作られるエージェントの多様性や強さの詳細な評価が挙げられる。方法としては、特徴的な戦略を持つエージェントをいくつか想定し、それらに対してパラメータを適切にチューニングすることでどの程度勝てるようになるかを見るというものが考えられる。

参考文献

- [飯田 03] 飯田弘之, 松原仁, ゲーム情報学の動向, 情報処理, vol.44, no.9, pp.895-899, 2003.
- [J. Schaeffer 97] J. Schaeffer, "One Jump Ahead: Challenging Human Supremacy in Checkers," Springer-Verlag, New York, 1997.
- [M. Campbell 02] M. Campbell, A. Hoane, and F. Hsu, "Deep Blue," Artif. Intell., pp.57-83, 2002.
- [D. Silver 16] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of Go with deep neural networks and tree search," Nature 529, pp.484-489, 2016.
- [片上 15] 片上大輔, 鳥海不二夫, 大澤博隆, 稲葉通将, 篠田孝祐, 松原仁, 人狼知能プロジェクト. 人工知能学会誌, vol.30, no.1, pp.65-73. 2015.
- [鳥海 17] 鳥海不二夫, 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐, 松原仁「ゲームプログラミング大会を用いた集合知的ゲーム AI 開発手法—人狼知能大会の開催と参加エージェントの分析—」, デジタルゲーム学研究, vol.9, no.1, 2017

*2 <http://aiwolf.org/control-panel/wp-content/uploads/2019/08/takeda.pdf>