所属陣営の仮定を利用した 3陣営人狼での投票行動学習の検討

植木 実結[1,a)]　横山 大作[1,b)]

概要：3 陣営人狼は敵陣営が複数あることから，盤面の状況次第でどこかの 2 つの陣営が結託して 1 つの陣営を倒す，敵陣営同士を争わせる，などというような 2 陣営よりも複雑な戦略が存在する．人狼知能の研究では 2 陣営配役の研究しか存在せず，そのような行動選択が難しい 3 陣営の研究はされていない．配役を 3 陣営に拡張させた人狼について，強化学習 (DQN) を用いて投票行動を学習させられるか検証したところ，行動の学習が難しい役職がいくつか存在した．これに対し，他のプレイヤの所属陣営の推定情報を明示的に与えることで学習が容易になるのではないかと考えた．本研究では，学習のうまくいかなかった役職について，個別の陣営推定ネットワークを用いて占い師真偽を仮定しその前提ば盤面から投票先を決定するエージェントと，投票行動学習と役職推定を同時に学習させたネットワークを用いたエージェントの 2 つについて勝率や投票率を比較・検証した．占い師真偽を仮定した状態での投票行動を行う手法では 2 役職とも勝率は向上しなかった．役職推定を併用して学習させた手法では併用しなかった場合より 1%ほど勝率が向上し，追放したい役職への投票率が上昇した．

A Study on Voting Action Learning in the Three-Team Werewolf Game Using Predictions About Team Affiliation

Ueki Miyu[1,a)]　Yokoyama Daisaku[1,b)]

Abstract: In a three-team Werewolf game, players can use more complex strategies than in a two-team game. For example, two teams can work together to defeat the third team, or players can encourage opposing teams to fight each other. These strategies depend on the game situation. Previous studies on Werewolf AI have focused only on two-team games. Three-team games, which require more complex decision-making, have not yet been studied. We investigated whether reinforcement learning using Deep Q-Networks (DQN) could teach agents to choose whom to vote for in a three-team game. The results showed that learning effective voting behavior was difficult for some roles. We hypothesized that explicitly providing estimates of other players' teams would make learning easier. For the roles that were difficult to train, we compared two approaches in terms of win rates and voting rates. The first approach used a separate team estimation network and selected voting targets based on an assumption about whether the seer was genuine. The second approach used a network that learned voting behavior and role estimation simultaneously. The first approach did not improve the win rate for either role. The second approach improved the win rate by about 1 percentage point compared with learning voting behavior alone. It also increased the rate of voting for the role the agents aimed to eliminate.

1. はじめに

人狼ゲームは多人数ゲームの 1 つであり，話し合いと投票によって追放を行う昼と，人狼が襲撃を行う夜を繰り返し，村人なら人狼をすべて追放すること，人狼なら自分が追放されないように村人を減らすことを目指す．3 陣営の人狼ではさらに妖狐が加わり，妖狐はどちらか陣営の決着がついたときに生存していれば勝利を奪える特性を持つ．3 陣営では，各陣営がそれぞれ盤面状況を見て，一方の敵

[1] 明治大学 (Meiji Uniersity)
[a)] ce255016@meiji.ac.jp
[b)] dyokoyama@meiji.ac.jp

© 2026 Information Processing Society of Japan

陣営を利用してもう一方の敵陣営を倒そうとしたり，一方の敵陣営と交渉したり，敵陣営同士を争わせるように仕向けたりするなどというような 2 陣営よりも複雑で幅広い戦略がある．

我々は以前に 3 陣営配役でルールベースエージェント 7 体に対して 1 体のエージェントについて DQN を用いて学習対象の役職を固定し，盤面情報がある程度そろってきた 2 日目以降の投票行動を学習した [4]．その結果，ルールベースエージェントより勝率が 2%ほど増加し，より追放したい対象の役職に投票できた．

この研究を発展させて学習対象をすべての役職に拡張し，同じくルールベースエージェント 7 体に対して 1 日目から投票行動を DQN で学習したところ，村人，占い師，背徳者の 3 役職でルールベースエージェントより勝率が下がった．特に村人ではほぼ学習ができておらず，味方の重要役職である占い師に対して投票をしてしまう行動が何度も見られた．そのため，投票行動学習において対戦相手の陣営を推定し，その推定結果を前提にした状態を入力として追加した場合に学習がうまくできるのではないかという仮説を立てた．

本研究ではこの仮説を検証する．勝率が下がった 2 役職 (村人，背徳者) について，提案手法 1 では占い師自称者たちの陣営推定を行い，その推定結果をもとにした行動出力を試みた．また，提案手法 2 では投票行動と役職推定を同時に行うネットワークを構築して，投票行動選択に影響があるかを調査した．

提案手法 1 は，DQN の途中出力を用いて新たに陣営推定を行うネットワークを構成する．構成したネットワークで陣営推定を行い，占い師自称者のうち村人陣営の可能性が最も高いエージェントを真の占い師と仮定する．その占い師が真である前提で盤面情報を更新し，その情報を DQN の入力として投票先を選択するエージェントを作成した．占い師の真偽仮定をしなかったエージェントと勝率を比較したところどちらも勝率向上にはつながらなかった．

提案手法 2 では，出力層を投票行動選択と役職推定の 2 つ用意して，同時に更新を行うネットワークを構築し，そのネットワークに対してルールベースエージェント 7 体を対戦相手として学習させ，役職推定と投票行動の評価を行う．背徳者・村人ともに，役職推定精度は占い師や人狼の推定精度は 71%以上，背徳者の占い師推定精度は 99%を超えた．一方，村人のそれ以外の 3 役職への推定精度は 56%を下回り，特に村人の妖狐推定精度は 35%と最も低かった．しかし，投票行動では村人の役職別投票率は妖狐が 5 役職で最も高い結果になった．勝率はもとの DQN の場合に比べて 2 役職とも 1%ほど上昇した．村人での勝率はルールベースエージェントの勝率を超えたが，背徳者はルールベースエージェントの勝率には届かなかった．

提案手法 1 では勝率は上がらず，提案手法 2 ではルールベースエージェントと比較して一部の役職でのみ勝率が上がるという結果になった．提案手法 2 での投票行動の評価では役職推定の出力結果自体を用いた投票決定は行っていないため，役職推定結果と用いて味方への投票を避けるような行動をとらせた場合に特に村人での勝率が向上するのではないかと推定される．今後は強化学習エージェントを対戦相手とした場合の勝率の変化や役職推定結果を用いた投票行動選択をするようなエージェントの構築・評価をしたい．

表 1 陣営の勝利条件

陣営	詳細
村人	妖狐が死亡したのちに人狼が全滅する.
人狼	妖狐が死亡したのちに人間の数が人狼の数以下になる.
妖狐	妖狐が生存している状態で，人狼が全滅するか人間の数が人狼の数以下になるかのどちらかを満たす.

表 2 役職

役職名	陣営	詳細
村人	村人	能力はない.
占い師	村人	夜に選んだ 1 人の対象が人狼かどうか知ることができる.
人狼	人狼	夜に選んだ 1 人を襲撃できる.
妖狐	妖狐	人狼に襲撃されても死亡しないが占い師に占われると死亡する (呪殺).
背徳者	妖狐	妖狐が誰か知ることができ, 妖狐が死亡すると道連れで死亡する.

2. 配役設定

本研究では以前の研究に引き続き，村人 3，占い師 1，妖狐 1，背徳者 2，人狼 1 の 8 人の配役を用いる．各陣営の勝利条件を表 1 に，各役職の説明を表 2 に示す．

妖狐が生存状態で先に人狼が死亡した場合，その時点で妖狐が勝利し，村人と人狼は敗北する．そのため，村人陣営は人狼よりも先に妖狐を倒す必要があり，逆に妖狐が生存しているときは人狼らしい対象を追放してはいけない．また，人狼も自身が追放されないように立ち回るだけでなく，妖狐が誰なのか探して倒す必要がある．

3 陣営ではある 1 つの陣営の勝利できる確率がすでに 0%である盤面が存在する．例えば村人，妖狐，背徳者，人狼の 4 人の盤面のとき，村人は誰が追放されても勝利できない．その陣営の投票によって残り 2 陣営の勝敗が確定することを防ぐために 3 陣営では引き分けが存在する．ここでは最終日となる 4 人盤面で投票が 2 回連続で同票で，投票された対象に妖狐が含まれた場合に引き分けになる．

3. 関連研究

人狼をプレイするようなエージェントの構築について，一般的な 2 陣営配役 (村人と人狼のみ) での陣営・役職推定を行った研究はいくつか存在する．SVM を用いた推定 [1]

© 2026 Information Processing Society of Japan

では 15 人人狼について人狼推定を行い，推定結果の最も高いエージェントへ投票するようにしたところ，勝率が推定していないエージェントより 1.7%ほど上昇した．ニューラルネットワークを用いた推定では [2] では 2017 年の人狼知能大会の 5 人人狼のログをもとに占い結果，投票数，生死，肯定・否定意見の対象を入力とした役職推定とそれと別に人狼に特化した推定の学習を行い，70%程度の推定精度を出し，大会の上位エージェントと対戦させて上位エージェントと同程度の勝率を出した．役職推定と DQN(Deep Q Network) での強化学習の併用 [3] では 15 人人狼で盤面情報と発話の方向性の種類を入力として学習をした．役職推定では 98%以上の推定精度を出し，人狼陣営では高い勝率を出した．

2 陣営では高い役職推定が行えることが検証されているが，3 陣営ではある 1 つの陣営を倒すために残り 2 つの陣営が同じ役職を狙って投票を行うなどというような行動もあり得る．そのため陣営推定や役職推定を利用することで勝率にどのような影響があるか，推定精度の上昇が勝率上昇につながるのかを検証する．

4. 予備評価：すべての役職について DQN エージェントを構築

4.1 実験の設定

学習対象を 1 体に固定し，以前に我々が作成した [4] ルールベースエージェント 7 体を対戦相手として 1 日目からの投票行動を DQN を用いて学習させる．学習条件として占い師は必ず 1 日目に CO を行うという条件のみを与える．DQN の更新は 343 ゲームごとに行い，報酬は勝利時に+1，敗北時に-1，引き分け時に 0 を与える．学習対象を村人，占い師，妖狐，背徳者，人狼の各 5 役職にそれぞれ固定して 117649 ゲームずつ学習を行う．

以降，この学習対象のエージェントのことを「自身」と表現する．

4.2 行動モデル

DQN の入力状態については，自身が死亡しても自身の勝敗が確定していない場合は 1 つの状態として扱う．自身が死亡していた場合投票行動を行うことはできないが，勝敗が決まっていないときには状態として定義する必要がある理由は，例えば 1 日目に決着がついて勝利した場合と，1 日目に自身が死亡し 3 日目まで他のエージェント同士で対戦が続いた末に勝利した場合を同じ状態で行動したときの結果として扱うことは不適切であるためである．

7 体の対戦相手エージェントについて，それぞれ盤面情報を以下の 14 個のパラメータで表現し，これを DQN の入力の状態として扱う．

- p_0: ステータス値 (表 3)
- p_1: 生存状態値

表 3 各投票候補へのステータス値

条件	数値
真の役職視点で妖狐が確定している	+96
真の役職視点で人狼が確定している	+64
真の役職視点で人外が確定していない	+32
村人を CO (占い師を CO していない)	+16
真の可能性のある占い師から人狼判定をもらっている	+8
CO している役職視点で人外が確定している	+4
CO している役職視点で人間が確定している	+2
自身への (背徳者の場合は妖狐への) 投票意思がない	+1

- p_2: 得票数 (自身の投票先の得票を除く)
- p_3: 役職 CO 時の死亡人数
- p_4 - p_6: 各占い師自称者の占い結果情報 (3 人まで)
- p_7 - p_{13}: 各エージェントから投票された回数

p_0 のステータス値は表 3 の合計値で与える．ステータス値を one-hot ベクトルで定義しなかった理由は，one-hot で定義して実験を行ったところ，この手法よりも不適切な投票先を選択することが多くなり，勝率が落ちたためである．p_1 の生存状態値は，対象のエージェントが死亡していた場合は-1 を，自身と対象エージェントが両方生存している場合は 1 を，自身が死亡していて対象エージェントが生存している場合は 0 を与える．p_2 の得票数は再投票が行われた場合に初期投票で対象のエージェントに投票された票数を表す．p_3 の役職 CO 時の死亡人数については学習条件により 0 固定である．占い結果情報では，1 人目の占い師自称者の情報を p_4 に，2 人目の占い師自称者の情報を p_5 に，3 人目の占い師自称者の情報を p_6 に入れる．占い師自称者がいない場合はすべて 0 となる．対象エージェントが占い師自称者本人の場合は 2 を，人狼でないと判定を受けている場合は 1 を，人狼であると判定を受けている場合は-1 を，それ以外は 0 を与える．p_7〜p_{13} については，対戦相手のエージェントの並び順に 1 人目を p_7，2 人目を p_8... として，そのエージェントから投票された回数を与える．

入力は，7 体の対戦相手についてこの 14 個のパラメータを与えてつなげた 98 次元の値とする．中間層は 2 層で 128 次元とし，出力の次元数は自身を除いたエージェント数の 7 とする．このうち出力値が最大の生存しているエージェントに対して $eps = 0.9$ の確率で投票行動を行う．残り 0.1 の確率でランダムに投票する．

エージェントの並び順はエージェントの固有 id に関係なく，ステータス値の昇順で定める．ステータス値が同じ値の場合はその中でランダムに順番が割り振られる．これはエージェントの固有 id という本来盤面情報として不適切な情報によってエージェントの並び順が決まることを防ぐためである．

4.3 評価

各 5 役職について，学習を行ったときの勝率を図 1 に，

© 2026 Information Processing Society of Japan

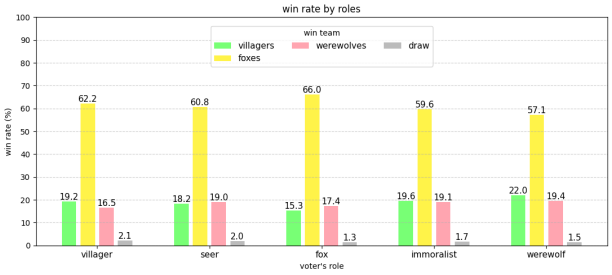


図 1　各学習対象役職別の勝率

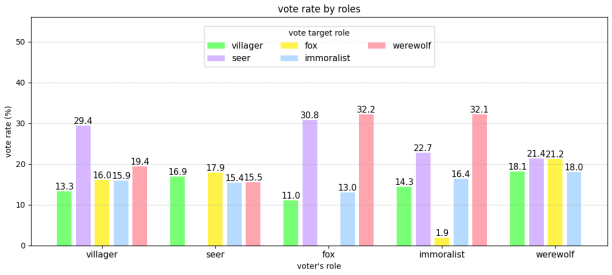


図 2　各学習対象役職別の投票率

表 4　ルールベース 8 体の対戦時の勝率

陣営	勝利数	勝率
村人陣営	23069	19.6%
妖狐陣営	74457	63.3%
人狼陣営	18115	15.4%
引き分け	2008	1.7%

各役職への投票率を図 2 に示す．横軸は学習対象自身の役職を示しており，villager は村人，seer は占い師，fox は妖狐，immoralist は背徳者，werewolf は人狼を示す．

それぞれルールベースエージェント 8 体での対戦と比較すると，妖狐と人狼については勝率が上がった一方で，村人，占い師，背徳者は勝率が落ちた．役職別の投票率の比較では，村人を除く 4 役職は，適切な投票対象へ投票できていたことも多い一方で，DQN の更新ごとに役職別の投票率が大きくばらついてしまう傾向があった．

村人については適切な投票位置を選択できず，味方の占い師や，その占い師の占った対象へ投票する行動が見られた．学習ログをいくつか取り出して確認したところ，村人がこのような不適切な投票行動をとる理由としていくつか考察ができた．

1. 元の勝率が低いことから村人として最適な行動をとっていたとしても敗北することが多く，学習が進まないこと．
2. 村人視点で占い師が 2 人以上 CO した場合にどちらが真の占い師か判別ができないことが多く，想定される役職候補の数を絞れていないこと．
3. 学習が進まないことでルールベースエージェントよりも適切な行動がとれないことから，学習対象自身が最後まで生存するよりも自身が早期に死亡して，ゲーム中盤以降の行動を味方のルールベースエージェントに任せるほうが勝率があがると判断し，あえて味方の占い師などに投票を入れることで再投票時に他エージェントから怪しまれて投票されて 1 日目に自ら追放される方針をとっている可能性があること．

5. 提案手法 1: 陣営推定 NN を構築し，その推定結果から占い師真偽を仮定した投票行動決定の検証

村人は最も所持している盤面情報量が少なく，占い師が複数人 CO を行っている場合に占い師の真偽が判明しづらく，役職候補の組み合わせが多くなって学習が進まないのではないかと推定された．それぞれの占い師が真だったときに盤面がどうなるか場合分けして考えることが DQN ではできないため，投票対象が学習できていない可能性があると考えられる．

1 人の占い師自称者を真であると条件づけたとき，条件を付けなかった場合と比べて各エージェントの役職候補を大きく絞れる可能性が高い．そのため，占い師が複数人 CO した場合，1 人の占い師自称者を真であると置いた状態で投票選択を行う．その占い師が真であることと，その占い師の占い結果を確定情報として含めたステータス値に変更した前提のもとで盤面情報を組みなおしたときの状態を DQN の入力として，その出力値によって投票選択を行う．

占い師の真偽の手がかりとして，各エージェントの陣営の推定を行う新たな 1 層のニューラルネットワーク（以降，陣営推定 NN と呼ぶ）を生成し学習させる．その学習させたネットワークを用いて，占い師自称者のうち最も村人陣営の可能性が高いエージェントを真の占い師と仮定した盤面情報を DQN の入力として投票行動を選択する．

5.1 陣営推定ネットワークの構成

投票行動を学習させた DQN の出力層の 1 つ手前の層 (128 次元) を，2 つのネットワークの最終段の共通入力として用いることにする．以降の説明では，この層を共有層と呼ぶことにする．陣営推定 NN では共有層の値を入力とし，各エージェントに対して各 3 陣営の可能性を示す値を出力 (21 次元) とする．教師あり学習を行い，正解の陣営を 1，不正解の陣営を 0 と置いたラベルを正解データとする．学習データは DQN の学習時のログからランダムに 16807 ゲーム取り出した．

5.2 ネットワークを利用するゲームプレイ手法

予備評価で学習させた DQN を陣営推定の利用と，占い師真偽を仮定した状態での投票行動の利用で合計 2 回用いる．1 回目の入力では占い師の真偽をつけずエージェントの自役職のみ定めた状態での入力を利用して，投票行動に用いる出力層の 1 層手前の共有層での値を出力させる．

陣営推定 NN は，その共有層の値を入力として，村人・

© 2026 Information Processing Society of Japan

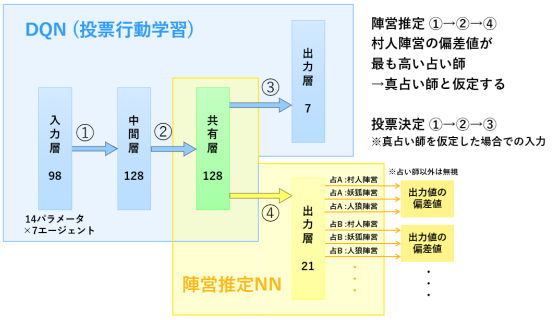


図 3 構成するネットワーク

妖狐・人狼のどの陣営に属するかを one-hot のベクトルで出力するよう学習する．この陣営推定 NN で出力された 21 次元の値を各エージェントそれぞれに対して 3 つの陣営ごとに分けて集計し，各陣営に対して偏差値を求める．陣営推定 NN を用いて出力した偏差値のうち，占い師自称者の中で最も村人陣営の偏差値が高かったエージェントを真の占い師であると置き，今度は DQN に真の占い師のエージェントを決め打った状態での入力値を与えて，そこでの出力値に対応したエージェントに投票する．

出力に対して偏差値で比較する理由として，ある占い師自称者 A_1 の村人陣営の出力値が他の占い師自称者 A_2,A_3... の出力値より高くても，A_1 の妖狐陣営や人狼陣営の出力値が A_1 の村人陣営の値よりさらに高ければ「村人陣営の可能性が高い」と判断できないからである．

5.3 陣営推定による占い師真偽判定を用いたエージェントの評価

予備評価で勝率の下がった村人と背徳者の 2 役職について 16807 ゲームずつ行い勝率を予備評価と比較する．

占い師の真偽判定の全体と占い師真偽未確定時のみでの正解率と，全体の勝率を 5 に示す．ただし，真偽判定の正解率は，各投票行動が行われるごとに占い師の真偽を正しく判断できたか，で示している．

真の占い師を陣営推定によって真の占い師であると正しく判定できた確率が村人が 64.6%, 背徳者が 79.7%，占い師の真偽が盤面上未確定の場合に絞ると，村人が 40.6%, 背徳者が 47.7%となった．これは占い師が 3 人以上 CO 場合も含む．真偽判定はそこまで精度は上がらない状態だった．また，2 役職とも全体の勝率の向上にはつながらなかった．

情報量の少ない 1 日目では陣営推定が行いにくいことと，占い師真偽判定を投票時にその都度行っていて，真偽判定が変化することで一貫した前提で投票先を選択できていないこと，真偽判断を間違った場合に大きく勝率が落ちてしまうからだと考えられる．

6. 提案手法 2: 役職推定と投票行動の両方を行う NN を用いた検証

提案手法 1 と同じく村人と背徳者の 2 役職について，陣営よりも細かく分割した役職推定と投票行動の学習を同時に行うネットワークを構成する．対戦相手はルールベースエージェント 7 体であり，学習での対戦回数は，背徳者が 117649 ゲーム，村人が 79919 ゲームとなっている．

表 5 占い師真偽判定別

	村人	背徳者
真偽判定正解数	28140	31695
真偽判定正解数 (未確定のみ)	10574	7374
真偽判定不正解数	15441	8083
真偽正解率	64.57%	79.68%
真偽未確定時正解率	40.65%	47.71%
勝率 (全体)	18.03%	58.98%

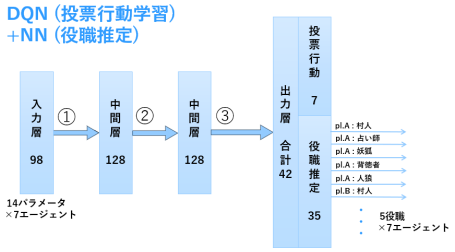


図 4 役職推定と投票行動を学習するネットワーク

6.1 ネットワークの構成

入力は予備評価と同じとし，中間層は 128 次元が 2 層，出力は投票行動に用いる 7 個に加えて，各 5 役職ごとの役職推定を行うための 35 個の，合計 42 個とする．(図 4)

予備評価と同じく 343 ゲームごとにネットワークの更新を行う．投票行動は強化学習で行い，役職推定は学習対戦ログとそのときの正解データを用いた教師あり学習で行う．誤差は，投票行動の誤差と，役職推定での誤差を 0.01 倍した値の合計値としてネットワークの更新をする．

役職推定の正解データは，各エージェントごとに，村人，占い師，妖狐，背徳者，人狼の 5 役職のうち正解の役職を 1，それ以外の役職を 0 と one-hot ベクトルでラベル付けしたものを，入力値の順番に対戦相手 7 体分並べた 35 個の値である．

6.2 ネットワークを利用するゲームプレイ手法

構成したネットワークを用いて対戦を行うときは，予備評価と同じ条件での入力値を与え，出力値の投票行動選択 7 つのうち最も出力値の高いエージェントへ投票する．役職推定での出力値はネットワーク更新のときの誤差として与える情報にのみ用いるため，投票行動選択では用いない．

6.3 役職推定の評価方法

役職推定の精度を検証・評価する．提案手法 2 での対戦ログはすべて役職推定の学習データとして使用されてい

© 2026 Information Processing Society of Japan

るため，テストデータは予備評価での対戦ログから 16807 ゲーム取り出したものを使用する．

提案手法 2 のネットワークの中の役職推定に使用される出力値 35 個について次の操作を行う．まず 35 個の値を 5 個ずつに分割し，各エージェント別の役職推定値 p_v, p_s, p_f, p_i, p_w に対して，各役職ごとにその役職の出力値とそれ以外の役職の出力値の差を合計した値をそれぞれ 5 役職に対して求める．この値を変換値とする．例えば，村人 p_v の場合の変換値 s_v は

$$s_v = (p_v - p_s) + (p_v - p_f) + (p_v - p_i) + (p_v - p_w)$$

となる．

それを対戦相手のすべてのエージェントに対して行った 35 個の変換値うち最大値のエージェントと役職の組み合わせを，このネットワークで判断した役職推定とする．また，残りの配役人数を考慮しながら，繰り返し残っている役職とエージェントに対応した変換値の集合のうちの最大値の組み合わせを順に役職推定の結果として定めていく．これを対戦相手のすべてのエージェントに対して行う．

各状態ごとの役職推定正解数 (最大値 7) の平均と，役職別での推定正解率で評価を行う．

6.4 役職推定の評価

役職推定の正解率をそれぞれ図 5(背徳者)，図 6(村人) に示す．total はすべての場合での役職推定の正解率を，seerCO は占い師を CO した場合での役職推定の正解率を示す．各状態ごとの役職推定正解数の平均は背徳者は 5.716 エージェント，村人は 4.061 エージェントだった．

図 5 より背徳者の場合では，99%以上の高い確率で真占い師のエージェントを推定でき，また 95%以上の確率で占い師を CO した人狼のエージェントを推定できていた．一方，占い師を CO しないエージェントのほうが人数が多いため，序盤の情報が少ないときに正確な推定を行うことが困難であり，村人，占い師を CO していない人狼やもう 1 人の背徳者の推定精度は占い師 CO のエージェントに比べて落ちた．全体で村人より人狼や背徳者の推定正解率が低いのは，真の村人のほうが自身を除く村人を騙る人狼・背徳者より人数が多いため，配役内訳の人数に合わせて役職推定を行えば少なくとも 1 人の村人は必ず村人と推定されるからである．

図 6 の村人の場合では，真占い師と人狼のエージェントはどちらも 71%以上の推定精度があり，占い師 CO を行っている場合は真占い師・人狼ともに 81%以上が正しく推定することができた．一方で妖狐推定の正解率は 35.6%と 5 役職で最も低かった．妖狐が占い師を CO していた場合の役職推定の正解率はわずか 3%ほどしかないのは，そもそも妖狐が占い師を CO することが人狼や背徳者と比べて極端に少ないからだと推測される．

提案手法 1 での占い師真偽の正解率と，提案手法 2 の真

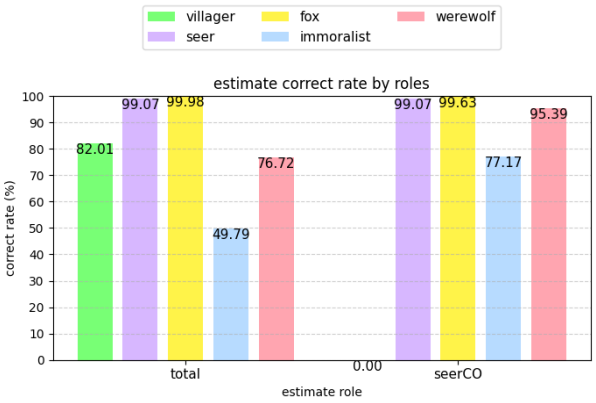


図 5 役職推定の正解率 (背徳者)

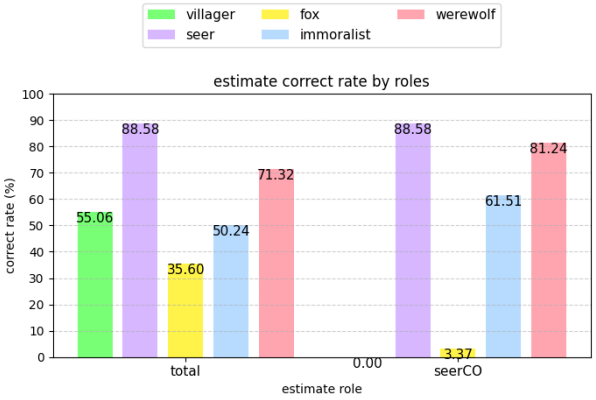


図 6 役職推定の正解率 (村人)

占い師の役職推定の正解率を比較すると，どちらも提案手法 2 のほうが大幅に精度上昇した．

6.5 投票行動の評価

投票学習に役職推定を同時に含めてネットワークの更新を行った場合に投票行動学習にどのような影響があるのか予備評価を比較対象として評価する．そのため，役職推定の出力結果自体を用いて投票行動決定を行ってはいない．

学習対象が背徳者の場合の役職別の投票率と勝率はそれぞれ図 7 と図 9，村人の場合の役職別の投票率と勝率はそれぞれ図 8 と図 10 のようになった．DQN only が予備評価の結果を，with Role estimate が提案手法 2 を示す．また，図 11〜図 14 は学習時の投票率の変化をネットワークが更新される 343 ゲームごとに示したものである．横軸は，次のネットワーク更新回数を示す．

勝率は，背徳者・村人ともにそれぞれ提案手法 2 のほうが背徳者は 0.94%, 村人は 1.09%ほど上昇した．

背徳者の役職別の投票率 (図 7) は予備評価よりも人狼への投票率が 7%以上増加した．背徳者視点では，人狼を追放できたその時点で勝利が確定するため，人狼が最も追放優先度の高い役職である．そのため，予備評価よりも提案手法 2 のほうが適切な投票先を学習できたと推定される．

© 2026 Information Processing Society of Japan

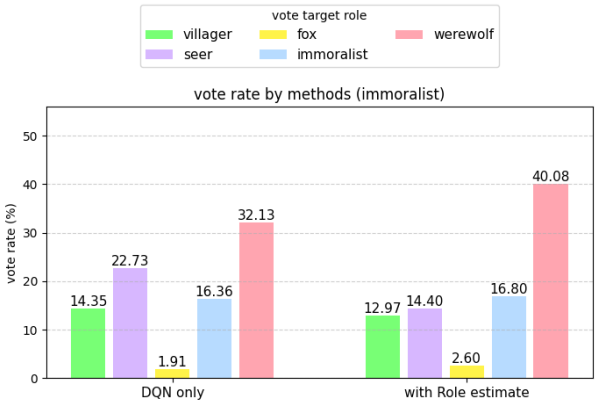


図 7 役職別投票率 (背徳者)

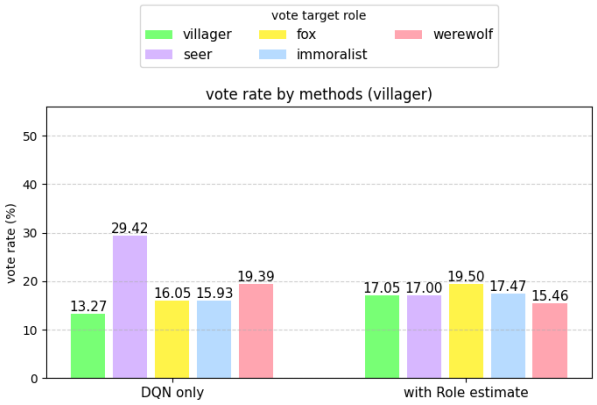


図 8 役職別投票率 (村人)

村人の役職別の投票率 (図 8) では，予備評価で味方の重要役職である占い師への投票率が高いことが最も問題であったが，提案手法 2 では占い師への投票率は 12%ほど下がった．しかし，89%ほどで味方である真占い師のエージェントが正しく推定できているが，その占い師への投票率は村人から見て味方なのか判断がつきにくい他の役職の投票率とそこまで差がなく，適切な投票先を学習できていない状態だった．一方，役職推定の評価では村人の妖狐推定正解率が 5 役職で最も低かったのにも関わらず，投票率では 19.5%と 5 役職のうち最も高くなったことから，正確に妖狐と推定できなかったとしても，妖狐の可能性のある対象を投票先として選択できる場合があると推定される．

予備評価のときの背徳者と村人の学習時の投票率の変化 (図 11, 図 13) はどちらも占い師への投票率が DQN の更新ごとに大きく変動し，学習がうまくできなかった．一方提案手法 2 では役職推定と投票行動を同時に学習させたことでネットワークの更新によって大きく投票率が変動することなく，予備評価のときよりも学習がうまくいった．また，役職の推定精度が必ずしも最適な投票選択につながる可能性が高いわけではないことが分かった．

予備実験で村人の投票行動が学習できなかった理由として，最適な行動をとっていても元の勝率が低いことから敗北することが多い可能性があると推定した．一方で，役職推定を用いることで味方を推定し，味方の可能性が高い対象に対しての投票を避けるような投票行動を行わせることで，元の勝率が低かったとしても妥当な投票先を選択することができる可能性が高く，さらに村人の占い師への投票率を下げることが可能であると推定される．

7. まとめ

3 陣営人狼において強化学習を用いた自陣営に利益になる行動をとれるエージェントを構築するために以前に特定の役職について 2 日目からの投票行動を DQN で学習させた．1 日目からに拡張して学習させたところ，一部の役職で勝率が落ち，特に村人では味方の占い師への投票行動が

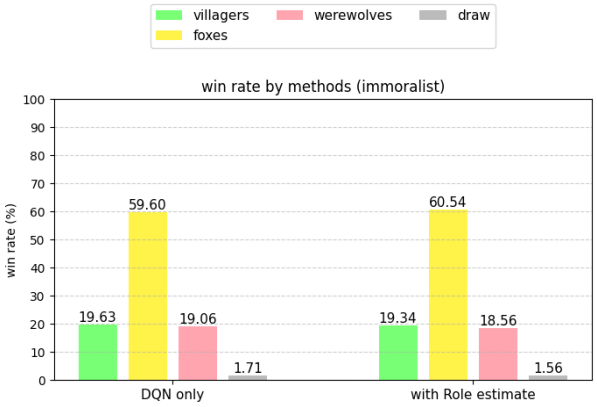


図 9 手法別の勝率 (背徳者)

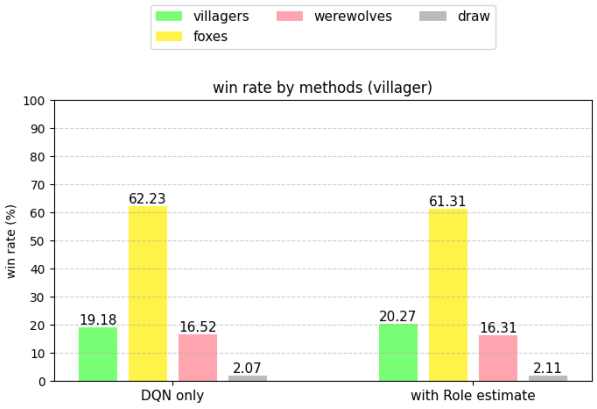


図 10 手法別の勝率 (村人)

多く学習がうまくいかなかった．その原因として想定される役職候補の数が多すぎることが推定された．そこで盤面の想定数を絞るために，陣営や役職の推定によって投票行動にどれくらい影響するかを 2 つの手法で検証した．

提案手法 1 での陣営推定 NN を構築し，その推定結果から占い師真偽を仮定した投票行動決定の検証では，真の占い師を正しく真であると推定できた確率はそこまで高くなく，勝率の向上にもつながらなかった．

提案手法 2 での役職推定と投票行動の両方を同時に学習させるネットワークでは，真占い師の推定精度は背徳者・村人ともに提案手法 1 よりも大きく上昇した．一方で占い

© 2026 Information Processing Society of Japan

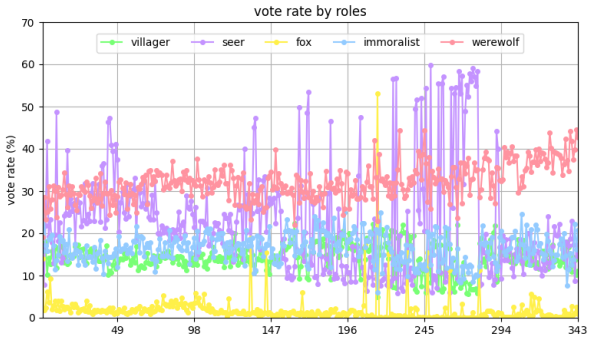


図 11 学習時の役職別投票率の変化 (背徳者/予備評価)

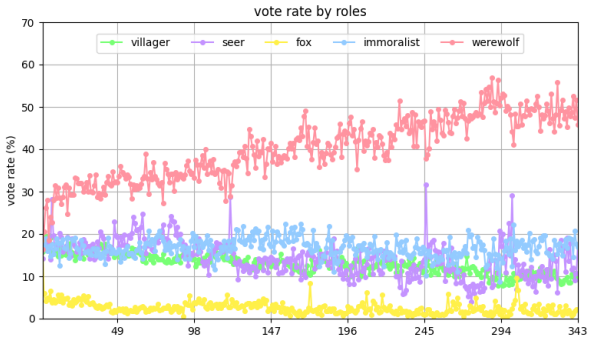


図 12 学習時の役職別投票率の変化 (背徳者/提案手法 2)

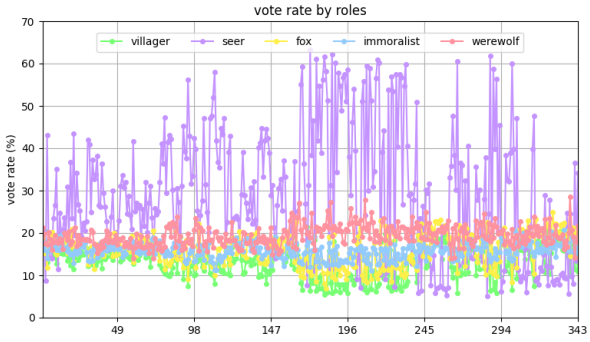


図 13 学習時の役職別投票率の変化 (村人/予備評価)

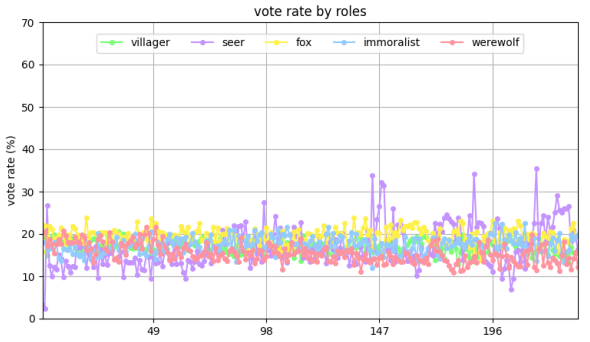


図 14 学習時の役職別投票率の変化 (村人/提案手法 2)

師を CO していない対象の役職推定はそこまで高くない結果になった．村人の妖狐推定精度は 5 役職で最も低かったが，提案手法 2 での村人の妖狐への投票率は上昇した．また，投票行動の評価では予備評価のときよりも勝率が背徳者・村人ともに約 1%上昇した．これにより，役職推定の精度の向上自体が最適な投票行動選択につながる可能性が高いわけではないが，役職推定を同時に学習させたことでより妥当な投票行動対象を選択することができるようになった．しかし，村人の占い師への投票率は予備評価よりも大きく改善されたものの，それでも投票率は高かった．

提案手法 2 では役職推定の出力結果を投票行動決定に使用することはなかったため，村人に対しては役職推定の出力結果をを用いて味方の可能性の高いエージェントを投票先から除外するような投票設計を行うことで村人の勝率はさらに上昇するのではないかと推定される．また，役職推定を利用することで 3 陣営特有の戦略である，どちらの敵陣営を狙うか，どちらの敵陣営の行動を利用するかなどの方針に対して投票行動を決めることができると期待される．

今後は，対戦相手をルールベースエージェントだけでなく強化学習エージェントに変更した場合にルールベースエージェントと同じように自陣営の勝利につながるような投票先を選択できるか調査するとともに，提案手法 2 のネットワークを用いて役職推定結果を投票行動決定に対してどのように用いるかをいくつか設計し，勝率や投票率を検証したい．

参考文献

[1] 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

[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] 福田 宗理, 穴田 一. 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. 2022 年度電気・情報関係学会九州支部連合大会, 2022

[4] 植木 実結, 横山 大作. 3 陣営人狼における Deep Q Network を用いたエージェント構築の検討. 第 30 回ゲームプログラミングワークショップ, 2025

© 2026 Information Processing Society of Japan