

進化論的アプローチによる 人狼ゲームにおけるエージェントの構築に関する調査

関井郁弥
横浜国立大学
sekii@forest.eis.ynu.ac.jp

阪本浩太郎
横浜国立大学
sakamoto@forest.eis.ynu.ac.jp

渋谷英潔
国立情報学研究所
shib@nii.ac.jp

森辰則
横浜国立大学
mori@forest.eis.ynu.ac.jp

1 はじめに

近年、スマートスピーカーの普及によってコンピュータとの対話が身近なものとなり、今後は情報検索や家電操作以外にも対話技術が応用されていくことが予想される。情報検索や家電操作などではユーザの指示に忠実に従うだけで問題ないが、例えば、商品セールスなどに应用することを考えると、ユーザが当初目的としていなかった商品売り込むといったユーザの意思決定に大きく影響を与える対話が求められることになる。我々は、ユーザの意思決定に大きく影響を与える対話を**翻意誘導対話**と定義する。しかしながら、翻意誘導対話をどのように実現するかに関する研究はほとんどない。我々は、自分の目的を達成するために相手の意思決定に影響を与えるという点で説得も欺瞞も翻意誘導対話の一種¹⁾と考えている。そこで、ゲームにおける勝敗のように目的が明確であり、かつ、その目的を対話によって達成するものが、翻意誘導対話の対象として適切であると考え、人狼ゲームを対象として研究に取り組むこととした。

人狼ゲームとは、対話を通して「村人」の中に潜伏した「人狼」を見つけ出す対戦型の多人数ゲームであり²⁾、近年、人狼知能プロジェクト³⁾など研究テーマとしても注目されている。人狼知能プロジェクトでは、「人間と自然なコミュニケーションを取りながら人狼をプレイするエージェント(AI)の構築」を究極の目標としており、エージェント同士の対戦成績を一つの指標として人工知能の研究開発に取り組んでいる。また、我々は、これまで人狼ゲームプラットフォーム LiCOS[1]の開発や人狼エージェントの分析[2]などに取り組んでいる。人狼ゲームは意思決定の際に相手がどのような行動をとるか分からない不完全情報ゲームであり、推論に誤謬がなくとも正解を導き出せるとは限らず、ゲーム中の行動が必ずしも論理的に説明できるわけではない。そのため、エージェントの発言が翻意誘導できたかを評価するためには、何度もゲームを繰り返して定量的に見積もる必要があり、模索段階で人間を相手に

ゲームを行うことは効率的でない。また、翻意誘導された過程を調査する上でも、思考過程が透明であるエージェントを対象とすることが望ましい。以上の背景から、我々は人狼ゲームにおける翻意誘導の対象となるエージェントの構築に取り組む。将来的に他のプレイヤーの発言の理解や、自然言語を用いた発話といった対話機能を含めた推理・発言の機能を持つエージェントの構築を検討する。人間の思考過程を模した構築したエージェント相手への翻意誘導を観察し、人間に対しても同様に翻意誘導ができるかの検証につながる情報を得ることを期待する。

翻意誘導の対象となるエージェントは、誘導される知性を備えつつ、エージェントごとにある程度の多様性があることが望ましい。そのようなエージェントを構築する手段として、我々は進化論的アプローチをとることとした。人狼ゲームと同じ不完全情報ゲームである繰り返し囚人のジレンマにおいて進化論的アプローチにより最適戦略を求める研究は多くなされている[3, 4, 5, 6, 7]。しかしながら、ゲームの複雑さが異なるため、そのまま人狼ゲームに应用できるか自明ではない。それゆえ、本稿では、進化論的アプローチによるエージェント構築に関する設計を行う。

2 エージェントの設計

エージェントの思考過程を明示的に表現するためには、エージェントがそれぞれ対戦相手の他エージェントの役職をどのように推理しているかという認識を表現する必要がある。エージェントの多層的な認識(例:「エージェントCは『エージェントAがエージェントBの役職を村人だと思っている』と認識している」など)という思考を可視化して表現するために、エージェントがそれぞれ自身の思考を行う思考空間(メンタルスペース)と、参加している各プレイヤーに対応しメンタルスペース内に存在する仮想上のエージェント(メンタルエージェント)を所有しているモデルを導入する。

エージェントの動作は、他のプレイヤーからの発言を受けてエージェントがゲーム状況を把握するためのメンタルスペースの更新と、ゲームを進行させるためメンタルスペースに対応した発話や投票を行うといった行動の2つからなる。

1) 翻意誘導のための発言が話者の認識する事実と一致していれば説得、不一致であれば欺瞞と考えられる。

2) 人狼ゲームそのもののルールに関しては、Wikipediaのサイト(<https://ja.wikipedia.org/wiki/汝は人狼なりや>)などを参照されたい。

3) <http://aiwolf.org/>

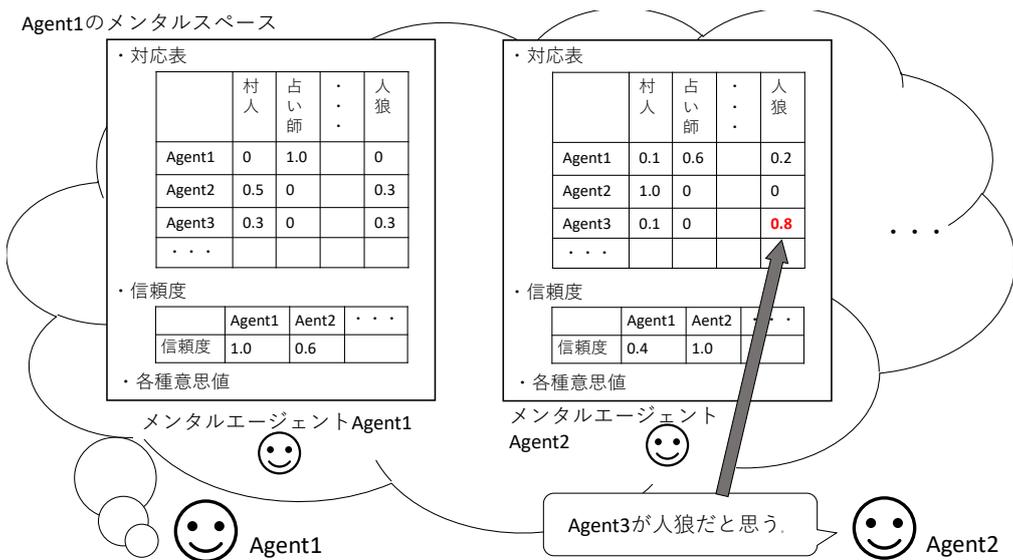


図1 メンタルスペースの構造

2.1 メンタルスペースの構造

図1にメンタルスペースの構造を示す。メンタルエージェント内には、そのエージェントが各エージェントの役職をどのように認識しているかというキャラクタと役職を対応させた表(対応表)を持つ。対応表は、プレイヤーと役職の表があり、対応はプレイヤーがそれぞれの役職である確率[0,1]として表される。エージェントのメンタルスペース内のメンタルエージェント1人につき、15人のプレイヤーごとに6つの役職それぞれの確率の対応表を所有していることで、「エージェントCは『エージェントAがエージェントBの役職を人狼だと思っている』と認識している」などという複雑な認識を扱うことができる。

また、エージェントはたとえ同じ内容の発言であっても発言者が異なる場合に、その発言の信用具合が変化することが考えられる。エージェント毎の信用具合の差を表すために、それぞれのメンタルエージェントには各プレイヤーの発言をどれほど信用するかという指標となる信頼度を所有させる。信頼度は各エージェントの発言を信用している度合い[0,1]として表される。信頼度が高いプレイヤーからの発言は対応表に大きな影響を与え、反対に低いプレイヤーからの発言は対応表に影響しにくくなる。信頼度も対応表と同様に、メンタルスペース内のメンタルエージェント1人につき、15人のプレイヤー全員への信頼度を所有させる。

誰に投票するか、人狼のプレイヤーが誰に襲撃するか、占い師や狩人のプレイヤーの能力の使用先を誰にするか、などのゲーム内での意思決定を含む行動を説明可能にするために、メンタルエージェントにはこれらの各種意思決定の際に用いる投票意思、占い意思、護衛意思、襲撃意思という意思値を所有させる。意思値は、あるプレイヤーに対して投票、人狼の襲撃、占い師の占い、狩人の護衛といった行動をしようという意思の強さを表した値で、基本的に意思値が最も高いプレイヤーに対して投票・襲撃といった各種行動を試みる。各種意思値はメンタルエージェン

トごとに1つずつ所有しており、対応表や信頼度の数値によって変動するほか、CO(カミングアウト:自らの役職を宣言する)などの発言内容にも応じて変化させる。

対応表、信頼度、各種意思値といった値は、「占い師COをがされたとき、発言者が占い師かもしれないと考える」「発言内容に矛盾が生じたとき、その発言者が人狼であると疑う」などといった、理性的なプレイヤーが人狼ゲームをプレイする際に用いるルールに従って変化する。これらを変化させる際に、それぞれのルールの持つ影響力の指標として、各エージェントにパラメータを持たせる。パラメータはそれぞれ村人陣営の役職のときに強く発現するパラメータ、人狼陣営のときに強く発現するパラメータ、陣営にかかわらず発現するパラメータの3種類が存在する。表1にエージェントの持つパラメータの一部を記載する。実際のエージェントの構築ではこのようなパラメータを多数使用するが、ここでは一部のみ抜粋する。このパラメータの組み合わせにより、エージェントの個性となる推理方法や発言などの行動が変化する。本研究では、このパラメータを遺伝的アルゴリズムにおける学習対象の遺伝子とし、この数値を最適化することによって理性的に推理を行い、行動するエージェントの構成を試みる。

2.2 意思決定

エージェントは発言を要求されるたびに、

1. 役職が占い師、霊能者でそのゲーム中にまだCOしていないければCO発言
2. 占い師、霊能者のときは能力の使用結果を公表
3. 一番投票してほしいエージェント(投票意思が最も高いプレイヤー)へ投票するように希望する発言
4. 今自分が誰を人狼だと思っているかなどの役職考察発言

の優先度で発言を行う。このアルゴリズムは第2回人狼知能大会のチーム鯰の発言アルゴリズム[8]を参考にした。

また、エージェントは投票を要求されるたびに、

表1 エージェントの構築に用いるパラメータ例

パラメータ名	概要	取りうる値	発現する陣営
古い騙り確率 (人狼)	役職が人狼の時に、古い師に CO する確率	[0,1]	人狼陣営
考察発言閾値	自分の対応表において、値がある閾値以上のとき、考察発言をする	[0,1]	共通
古い師 CO の発言者	古い師 CO 発言がされたとき、発言者の対応表において、発言者の古い師値を [パラメータ] だけ上げる	[0,1]	共通
古い師値増加	自分の役職が人間で、自分に対して古い結果人狼の発言がされたとき、自分の対応表において、発言者の信頼度を [パラメータ] だけ下げる	[0,1]	村人陣営
人間の古い結果人狼の非発言者信頼度変化			

表2 試合ごとに加算される得点

村人陣営		得点
ゲームに勝利したら		+ P_{V1} pt
人狼に投票するたびに		+ P_{V2} pt
人狼陣営		得点
ゲームに勝利したら		+ P_{W1} pt
ゲーム終了時に生存していたら		+ P_{W2} pt

生存している自分以外のプレイヤーの中で最も投票意思が高いプレイヤーへの投票を行う。人狼のエージェントは、襲撃先の投票が要求されるたびに、エージェントは生存している人狼でないプレイヤーの中で最も襲撃意思の高いプレイヤーへ投票を行う。

3 エージェント構築の環境

翻意誘導される上で知性を持ったエージェントでありながら、異なる個性をもったエージェントが多数得られることを期待し、並列分散型遺伝的アルゴリズム [9](並列分散 GA) によってエージェントを構築する。エージェントの個体が持つパラメータの集まりを遺伝子としており、人狼ゲームをプレイする上で最適な遺伝子パラメータの組み合わせを学習することを目的とする。

個体の遺伝子に多様性を持たせるために村を $N_{village}$ 個用意する。また、勝敗に運要素が絡むため、信頼できる勝率を得るためにそれぞれの村で N_{game} 回ずつ対戦を行う。

適者生存を示すスコアは村人陣営としてのスコアと人狼陣営としてのスコアの2種類を用意し、役職が村人陣営(村人, 古い師, 霊能者, 狩人)で参加した対戦での結果を村人陣営としてのスコア, 人狼陣営(人狼, 狂人)の役職で参加した対戦での結果を人狼陣営のスコアとして計算することで、村人として優秀な個体, 人狼として優秀な個体を判別する。村人陣営, 人狼陣営はそれぞれ試合ごとに表2の点数を加算する。ここで、得点 P_{V1} と P_{V2} , P_{W1} と P_{W2} については、各プレイヤーは自身の陣営の勝利を何より優先させるべきという考えから、 $P_{V1} > P_{V2}$, $P_{W1} > P_{W2}$ となることが想定される。今回の実験では、 $P_{V1} = 10$, $P_{V2} = 1$, $P_{W1} = 10$, $P_{W2} = 5$ とした。

各陣営としてのスコアは、その陣営として参加したゲームあたりの平均得点となる。(一度も役職を引かなかった陣営のスコアは0)

1世代分の対戦が終わると、遺伝・交叉により、次世代の個体を生成する。今回の実験では、エリート保存戦略 [10] によって4個体、交叉によって11個体を生成する。まず、人狼としてのスコアの1位、2位と村人としてのスコア1位、2位の個体を選出する。村人としての1位、2位と人狼としての1位、2位の個体計4個体をそのまま残し、人狼陣営1位

表3 遺伝子の交叉方法

村人側	人狼側	共通	個体数
v1	v1	v1	1
v2	v2	v2	1
w1	w1	w1	1
w2	w2	w2	1
v1	w1	v1	3
v1	w1	w1	3
v1	w2	v1	1
v1	w2	w2	1
v2	w1	v2	1
v2	w1	w1	1
v2	w2	v2 or w2	1

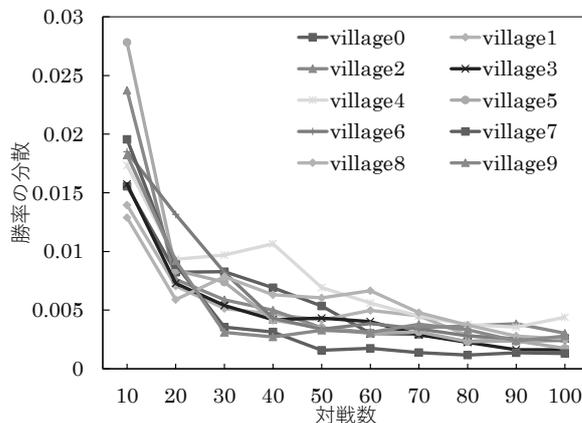


図2 同一エージェント同士の対戦の試合数と勝率

の個体(w1), 人狼陣営2位の個体(w2), 村人陣営1位の個体(v1), 村人陣営2位の個体(v2)の遺伝子をそれぞれに対して表3のように、村人としてのスコアの高かった個体から村人側の遺伝子を、人狼としてのスコアが高かった個体から人狼側の遺伝子を遺伝させる形で11個体を新しく生成する。共通して発現する遺伝子のv2 or w2の部分は、v2かw2のどちらかをランダムに遺伝させる。

この遺伝・交叉システムを用いて世代を重ね、学習の局所解への早期収束を防ぐため、 $N_{generation}$ 世代ごとに一度、遺伝・交叉の後に他の村への個体の移住を行った。個体の移住では、それぞれの村について村のエージェントのうちランダムに7個体を移住個体として選び、すべての村の移住個体を集めて、ランダムに各村に7体ずつ配りなおすという動作を行い、ほかの村のエージェントを交えて次世代の対戦を行うことで個体の多様性の維持を図る。

また、個体の移住とは別の手段での多様性の維持手段として、これらの遺伝・交叉・個体の移住が終了した後で、特定の変異率による突然変異を行う。今回の実験では、新しい世代の全個体に対してそれぞれ変異率0.01で遺伝子のうち1箇所をランダム

表4 各世代交叉終了時に最も数の多かった遺伝子の個体名と個体数

村		1 世代	2 世代	3 世代	4 世代
1	個体名	1007	1007	2001,2002	2002
	個体数	8	9	5	9
2	個体名	2015, 2016	2017	2017	2017
	個体数	3	10	14	15
3	個体名	2045,2046	2049	2049	2049
	個体数	3	4	14	15
4	個体名	2045,2046	2049	2049	2049
	個体数	3	9	11	14
5	個体名	2060,2061	2063	2065	2063
	個体数	3	10	7	11
6	個体名	2075,2076	3076	2077	2077
	個体数	3	4	13	15
7	個体名	1090,1091	2090	2092	2092
	個体数	3	5	13	15
8	個体名	1118	2108	1118	2108
	個体数	8	9	9	15
9	個体名	2120,2121	2120	2120	2120,2121
	個体数	3	8	12	7
10	個体名	2135,2136	2137	2137	2137
	個体数	3	9	15	15
11	個体名	2150,2151	2150	2150	2150
	個体数	3	6	15	15
12	個体名	2165,2166	2165	2165	2165
	個体数	3	8	9	8

に初期化する突然変異を行った。

4 遺伝的アルゴリズムの設定値

今回の実験における各種設定値について、一部の設定値について妥当な値を調査するため実験を行った。

4.1 ゲームルール

それぞれの村は 15 人狼を行い、役職の内訳は占い師 1, 霊能者 1, 狩人 1, 狂人 1, 人狼 3, 村人 8 人ずつの 15 人からなる。他、詳細なルールは人狼 BBS の G 国⁴⁾における 16 人村(初日の犠牲者がいないため実質同じルールとして働く)を参考にした。

4.2 1 世代あたりの対戦数について

自己対戦において 1 世代の対戦数に必要な対戦数を調査するため、同一エージェント同士の対戦を行った。同一エージェント同士の対戦では、まったく同じパラメータの組み合わせをもつエージェント同士による対戦を無限回に繰り返すことで理論上エージェントの勝率の分散はどの村でも 0 に収束するという考えのもと、同一エージェント同士の対戦 100 回を 10 セット行い、勝率の分散値の変化を調査した。

図 2 は、まったく同じ遺伝子の組み合わせをもつエージェント同士による対戦を 100 回 10 セットを行ったときの、試合数に対するエージェントの勝率の分散である。対戦数が 10 対戦付近ではまだ村ごとの分散が大きく、結果がばらけてしまっているが、30 対戦から 50 対戦で分散が小さくなり、そのあと分散は小さくなり続けているものの概ね横ばい

4) <http://ninjinix.x0.com/wolfg/>

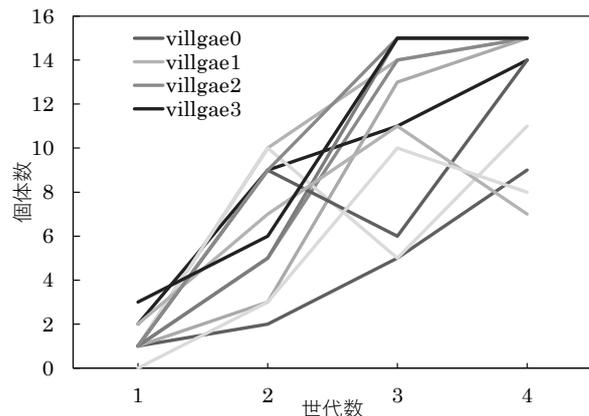


図3 第4世代終了時に最も数の多かった遺伝子の各世代での個体数変化

の値をとっている。この結果より、エージェント同士の勝率がある程度収束し、個体ごとに優秀な個体とそうでない個体を判別するのに対戦数は最低限 30 対戦から 50 対戦は必要となることが考えられる。そこで、本研究では 1 世代あたりの対戦数 N_{game} を 50 対戦として設定し実験を行った。

4.3 並列分散 GA の移住までの世代数

並列分散型遺伝的アルゴリズムによる、移住までに必要とされる世代数について最適な世代数を調べるための調査を行った。表 4 は、1 世代あたりの対戦数を 50 としたとき 12 の村について、各世代後の交叉終了時に最も数が多かった遺伝子の個体名と個体数を示したものである。遺伝子の個体名については、上 1 桁はその個体がどの世代で初めて生成されたものであるかを示しており、下 3 桁はその世代で何番目に生成された個体であることを示している。

表 4 の結果について、第 4 世代終了時に最多個体数である個体はほとんどが第 3 世代終了時でも最多個体であり、第 3 世代終了時点で過半数を占めているものが多いため、第 3 世代終了時の時点で交叉によって優秀な遺伝子はほとんど完成している可能性を考え、さらに第 4 世代終了時に個体数が最も多かった個体の各世代終了時の個体数について調査した。図 3 は、1 世代あたりの対戦数を 50 としたときの第 4 世代終了後の交叉終了時に最も数が多かった遺伝子の各世代での個体数の変化を示したものである。この結果より、第 3 世代から第 4 世代にかけて、最多数個体の個体数は増えているものの、第 3 世代までの結果でどの個体かその後最多数となるかについて大勢は決しているものと判断し、本研究では個体の移住までの世代数 $N_{generation}$ を 3 世代ごとと設定した。

5 まとめ

本稿では、翻意誘導の対象となりえる人狼エージェントの構築のため、進化論的アプローチによるエージェント構築に関する設計と、方法の調査を行った。今後は発話の意味理解や自然言語による発話の生成を含め、実際にエージェントを構築しエージェントに対しての翻意誘導を実現する技術の調査を行う予定である。

参考文献

- [1]阪本浩太郎, 永山翔滋, 石下円香, 渋木英潔, 森辰則, and 神門典子. 人狼ゲームプラットフォーム licos を用いた欺瞞対話コーパスのためのゲームログの収集. *言語処理学会第 25 回年次大会*, 3 2019.
- [2]Shoji Nagayama, Jotaro Abe, Kosuke Oya, Kotaro Sakamoto, Hideyuki Shibuki, Tatsunori Mori, and Noriko Kando. Strategies for an autonomous agent playing the “werewolf game” as a stealth werewolf. In *Proceedings of the 1st International Workshop of AI Werewolf and Dialog System (AIWolf-Dial2019)*, pages 20–24, Tokyo, Japan, October 2019. Association for Computational Linguistics.
- [3]Robert Axelrod. *Genetic Algorithms and Simulated Annealing*, chapter The Evolution of Strategies in the Iterated Prisoner’s Dilemma, pages 32–41. Morgan Kaufman, Los Altos, CA, 1987.
- [4]Richard Brunauer, Andreas Löcker, Helmut A. Mayer, Gerhard Mitterlechner, and Hannes Payer. Evolution of iterated prisoner’s dilemma strategies with different history lengths in static and cultural environments. In Yookun Cho, Roger L. Wainwright, Hisham Haddad, Sung Y. Shin, and Yong Wan Koo, editors, *Proceedings of the 22nd Symposium on Applied Computing*, pages 720–727. ACM, 2007.
- [5]Mieko Tanaka-Yamawaki and Taku Murakami. *Effect of Reputation on the Formation of Cooperative Network of Prisoners*, pages 615–623. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [6]糸井良太 and 田中美栄子. 進化型繰り返し囚人のジレンマにおける最適戦略の探究. Technical Report 3, 鳥取大学大学院工学研究科エレクトロニクス専攻, 鳥取大学大学院工学研究科エレクトロニクス専攻, may 2012.
- [7]鳥居 拓馬, 日高 昇平, and 真隅 暁. 学習あり繰り返し囚人のジレンマにおける協調行動の発生. *人工知能学会全国大会論文集*, JSAI2014:4H13–4H13, 2014.
- [8]狩野芳伸, 大槻恭士, 園田亜斗夢, 中田洋平, 箕輪峻, and 鳥海不二夫. *人狼知能で学ぶ AI プログラミング: 欺瞞・推理・会話で不完全情報ゲームを戦う人工知能の作り方*. マイナビ出版, 2017.
- [9]三木光範, 廣安知之, 畠中一幸, and 吉田純一. 並列分散遺伝的アルゴリズムの有効性. *日本計算工学会論文集*, 2000:20000038–20000038, 2000.
- [10]北野宏明. 遺伝的アルゴリズム. *人工知能学会誌*, 7(1):26–37, 1992.