

特集 「エンターテインメントにおける AI」

人狼知能プロジェクト

Project AI Wolf

片上 大輔

Daisuke Katagami

東京工芸大学工学部コンピュータ応用学科

Department of Applied Computer Science, Faculty of Engineering, Tokyo Polytechnic University.

katagami@cs.t-kougei.ac.jp, <http://www.hss.cs.t-kougei.ac.jp/>

鳥海 不二夫

Fujio Toriumi

東京大学大学院工学系研究科

Department of Applied Computer Science, Faculty of Engineering, The University of Tokyo.

tori@sys.t.u-tokyo.ac.jp

大澤 博隆

Hiroataka Osawa

筑波大学システム情報系

Faculty of Engineering, Information and Systems, University of Tsukuba.

osawa@iit.tsukuba.ac.jp, <http://hiroosa.com/>

稲葉 通将

Michimasa Inaba

広島市立大学大学院情報科学研究科

Graduate School of Information Sciences, Hiroshima City University.

inaba@hiroshima-cu.ac.jp

篠田 孝祐

Kosuke Shinoda

電気通信大学大学院情報システム学研究科

Graduate School of Information Systems, The University of Electro-Communications.

kosuke.shinoda@is.uec.ac.jp

松原 仁

Hitoshi Matsubara

公立はこだて未来大学システム情報科学部

Faculty of Systems Information Science, Future University Hakodate.

matsubar@fun.ac.jp

Keywords: werewolf AI, werewolf, communication game, general artificial intelligence.

1. はじめに

「ここは、とある村。平和に暮らしていた村人達にある日、衝撃的な事実が告げられる。この村の住人の中に、人狼が紛れ込んでいる——。人狼は完璧に村人になりすまし、見かけから判断することができない。しかし、夜になるとその本性を現し、毎晩一人ずつ村人を襲っていく。これに対して村人達は、昼間の間に議論を行い、多数決によってもっともあやしいと思われる人物を毎日一人ずつ処刑することにした。夜には一人の村人が襲われ、昼には人狼か村人のどちらかが処刑される。これを繰り返して、人狼をすべて処刑することができれば、村人達は平和な村を取り戻して勝利。生き残っている人狼の数が村人と同数になれば、村は滅んで人狼の勝利となる。」以上が、コミュニケーションゲーム「人狼」の大まかな世界観である【人狼読本 13】。古くからヨーロッパでエンターテインメントとして行われてきたこのゲームが、ここ数年、さまざまなメディアを通して、日本でブームを起し始めている。著者らは、この人狼のもつ魅力に注目した各分野で活躍する研究者が結集し、2013 年人狼をプレイする人工知能を実現しようと試みる人狼知能プ

ロジェクト【人狼知能プロジェクト 13】を立ち上げた。

本稿では、エンターテインメントとしての人狼、AI 領域における人狼、そして教育における人狼など、さまざまな視点を通して、この人狼知能プロジェクトが目指す人狼について説明する。

2. 人狼とは

2.1 人狼とはどんなゲームか？

人狼ゲームは、アメリカのゲームメーカ Loony Labs. が 2001 年に発売したパーティーゲーム「汝は人狼なりや？」【汝は人狼なりや？ 01】およびその派生ゲームの総称である。多数の類似ゲームが世界中で市販され、世界中でプレイされている。日本においてもタブラの狼【タブラの狼 02】やうそつき人狼【うそつき人狼 13】など多数のゲームが販売されている。人狼をプレイする方法としては、前述したような市販のカードなどを使ってテーブルを囲んで行う対面型と、電子掲示板などの Web 上のアプリケーションを使って行う BBS タイプが存在する。

2.2 人狼ゲームの流れ

プレーヤーにはまずランダムに「役職」が割り当てられる。プレーヤーは役職によって、人間または人狼陣営にそれぞれ振り分けられ、各プレーヤーはチームの勝利を目指す。人間陣営の目標は人狼の全滅に、人狼陣営の目標は人間の人数を人狼の人数と同数以下にすることにあり、目標を達成した陣営の勝利となる。

各自の役職は本人以外には非公開であるため、自分以外の誰がどの役職かわからない。特に、人間側は誰が人狼かわからないため、会話の中から人狼を探し出すことが基本的な行動指針となる。一方、人狼陣営のプレーヤーは同じ人狼陣営のプレーヤーをゲーム開始時に知らされる。そのため、人狼陣営に所属するプレーヤーは互いに協力しながら、人間陣営に正体がばれないように行動することが基本的な行動指針となる。

ゲームは昼と夜の二つのフェーズからなる。昼のフェーズではすべてのプレーヤーによって、誰が人狼かを探し出すための議論が行われる。このとき、後述する各種能力をもった役職についているプレーヤーは当該能力によって知り得た情報を用いて、自分達の陣営が有利になるように議論を導くことになる。一定期間の議論の後、プレーヤー全員の投票によって、人狼と考えられる人物を処刑する。処刑されたプレーヤーはゲームから除外され、ゲーム終了まで参加することができない。

夜のフェーズでは、人狼陣営に所属するプレーヤーが人間陣営のプレーヤーを一人選び、襲撃する。襲撃されたプレーヤーは死亡者として扱われ、処刑されたプレーヤーと同様にゲームから除外される。また、各種能力をもった役職には、夜のフェーズに能力に応じた情報を与えられる。昼のフェーズと夜のフェーズを繰り返し、勝利陣営を決定する。

議論において、人間陣営に所属するプレーヤーは人狼の嘘を見破るかが最大のポイントとなる。また、能力をもつ役職に就いたプレーヤーは能力によって知り得た情報を使って他のプレーヤーを説得することがポイントとなる。一方、人狼陣営のプレーヤーは自分達が不利にならないように議論を誘導し、時には能力をもった役職であると偽り、議論を間違った方向へ誘導することなどが基本プレイとなる。

2.3 人狼ゲームにおける主な役職

2.1節で説明したように、人狼にはさまざまなスタイルや、バージョンが存在し、役職もそれによって大きく異なるが、人狼 BBS [人狼 BBS] のデータ分析から始まった本プロジェクトでは、主に以下の役職を基本役職として考える。

(1) 村人

人間陣営に所属する。特に何の能力ももたない。

(2) 占い師 (予言者)

人間陣営に所属する。夜のフェーズで占い結果として

指定した一人が「人狼であるか否か」を知ることができる。人間陣営においては最も重要な役職である。

(3) 霊媒師 (霊能者)

人間陣営に所属する。前日に処刑した人物が人狼かどうかを知ることができる。

(4) 狩人

人間陣営に所属する。夜のフェーズでプレーヤー一人を指定して、人狼の襲撃から守ることができる。狩人が守ろうとしたプレーヤーを人狼が襲撃した場合、その日は誰も死なないことになる。

(5) 人狼

人狼陣営に所属する。複数人狼がいる場合は、互いにコミュニケーションを取ることが可能である (対面の人狼の場合、夜のフェーズで目配せとジェスチャーによるコミュニケーションを行う)。夜のフェーズで任意の村人を襲撃することが可能。

(6) 狂人

人狼陣営に所属する。ただし、人狼からは誰が狂人であるかは把握されず、狂人からも誰が人狼なのかはわからない。それ以外は特殊能力がないただの村人と同様であるが、人狼陣営が勝利したときに勝利となるため、ひそかに人狼に協力をする。勝利人数のカウント時には人間陣営として数えるため、終盤では狂人が自ら処刑される行動によって勝利することも可能である。

2.4 人狼ゲームの基本戦略

§1 人間陣営の基本戦略

人間陣営の基本戦略は人狼陣営の嘘を見抜くことにある。人狼を探し出すには占い師と霊媒師による情報が重要である。しかしながら、多くの場合人狼陣営のプレーヤーが偽の占い師および霊媒師として名乗り出る (以下 CO : Coming Out) するため、誰が真の占い師、霊媒師であるかを見抜くことが必要となる。

例えば、占い師 CO を行った二人の占い師のうち一人が自分を人狼であると指摘した場合、その占い師は人狼陣営 (人狼もしくは狂人) に所属していることがわかる。このような情報を積み重ねていくことによって、各プレーヤーは誰が人狼かを絞り込んでいく。

また、人間陣営では占い師や霊媒師がどのタイミングで CO するかを全員の相談であらかじめ決めておき、人狼陣営のプレーヤーが人間陣営のプレーヤーを騙す要素が少なくなるようにプレイすることが多い。

§2 人間陣営役職もちの基本戦略

人間陣営で役職をもつプレーヤーは占い師、霊媒師、狩人である。

占い師は占いによって直接人狼を判定できるため、人間陣営で最も重要な役職であるといえる。したがって、占い師は夜の占いによってできるだけ早く人狼を見つけ出すことが必要である。しかしながら、もし人狼陣営に占い師であることが発覚すると襲撃の対象となるため、

占い師であることは隠しておくことが多い。ただし、狩人が残っていることがわかっていく序盤に占い師であることを **CO** することも多い。しかしながら、人狼陣営のプレイヤー（人狼または狂人）が「自分が占い師である」と嘘を付くことが多いため、その場合は自分が本物であるという信頼を勝ち取るため他のプレイヤーを説得することが必要となる。

霊媒師は、処刑されたプレイヤーが人狼であるかどうかを判断できるが、能動的なアクションは起こし難いため、多くの場合人狼が処刑されたときに名乗り出てその事実を告げることになる。ただし、人狼陣営のプレイヤーが霊媒師を騙った場合は対抗して名乗り出る必要もある。この能力によって二人の占い師のどちらが本物かを判断する役割になることも多い。

狩人は他のプレイヤーを襲撃から守ることができる重要な役職であるが、自分自身を守ることはできないため、自ら狩人であることを名乗ることは少ない。狩人であるとはばれないように真の占い師を守ることが重要である。また、襲撃から誰かを守った場合、守られたプレイヤーが人間陣営であることがわかるため、他のプレイヤーよりも情報が多くなる場合もある。その場合は、得られた情報を使って議論をうまく誘導する必要がある。リスクはあるが、占い師は守られているだろうという人狼の想定をもとに、他の襲撃されそうな人を守るという駆け引きをすることも可能である。

§3 人狼陣営の基本戦略

人狼陣営の基本戦略は人間陣営を騙して人狼陣営のプレイヤーが処刑されないように議論を誘導することにある。人狼陣営のプレイヤーは基本的に人間陣営にいるように振る舞う。ただし、占い師の情報によって正体が暴かれる危険があるため、占い師を見抜き襲撃する、あるいは占い結果を他のプレイヤーが信じないよう議論を誘導する必要がある。

そこで、多くの場合複数いる人狼陣営のプレイヤーのうち何人かが占い師や霊媒師であると **CO** する戦略（騙り）を採用することが多い。このとき、人狼が **CO** する場合もあれば、狂人が勝手に **CO** する場合もある。

人狼が占い師を騙った場合、人狼とそれ以外のプレイヤーがわかっているため、常に正しい占いを行うことができるため、疑われ難い。一方狂人が占い師を騙った場合は誰がどの役職かはわからないため、間違った占いを行ってしまう可能性もあるが、それによって狂人自身が処刑されても人狼陣営の勝利に貢献することになる。ケースとしては狂人が占い師を騙り、人間判定をしていくことが多いが、適切なタイミングで人狼判定をすることで、人狼が自分を狂人と認識してくれるというメリットもある。ただし霊媒師がいると、次のターンで、人狼判定を出された人が人狼ではないということがわかってしまうので注意する必要がある。

2・5 オンライン型人狼（短期人狼、BBS 人狼）

日本におけるオンライン上の人狼は、短期人狼（チャット人狼、**CGI** 人狼）と、電子掲示板で行われる長期人狼（**BBS** 人狼）がある。カードゲーム型人狼を模したチャット人狼は議論が行われる「昼」の時間と、占いや襲撃など能力が使用される「夜」の区別があり、「昼」の議論時間は数分である。一方、長期人狼は日本独自の発展を遂げた人狼ゲームであり、夜の時間がシステムによって行われるため存在せず、昼の議論のみでゲームが行われる。昼の議論は 24 時間行われるため、長期人狼と呼ばれる。

インターネット上で行われる人狼について最も記録が古いのは、2003 年に誕生したチャット形式の「汝は人狼なりや？」である [aduma 03]。この形式の人狼はオンライン上で行われる **CGI** 型と呼ばれるようになった。**CGI** 型の人狼は主に匿名掲示板から参加者を集める形で発展している。**CGI** 型の人狼はカードゲーム型の人狼をそのままオンライン上に模しており、カードゲームにおける「昼」でのプレイヤー同士の議論と、人狼が処刑を決める「夜」をチャット上で再現し人狼を行う。

CGI 型に近い形式として、Yahoo! ジャパンでは、Yahoo! ゲームにおいて Web ゲームとして人狼ゲームのサービス [人狼 Online14] を 2014 年秋から提供を始めた。このゲームは、対面型の人狼ゲームを Web ゲームとして実装したものであり、最大 18 人のプレイヤーが人狼を行う **ROOM** に入り、プレイヤーの人数に応じて一定時間に制限された昼と夜のチャット時間を通して投票などの行動を決定する。システムは主にゲームマスターの役割を行っている。そのため、一定人数が参加しなければゲームが始まらないこともある。このようなゲームにおいては、場の盛り上がりによりリピート率に影響を与えるため、人狼知能プログラムのような仮想アバターの利用が想定される余地が多にある。

一方 **BBS** 型の人狼は、カードゲーム型の人狼と類似しているが、インターネット上の掲示板の特色を生かし、夜の間に人狼同士が秘密の会話を行えるような仕組みを加えている。日本における **BBS** タイプの人狼のうち、最も盛んにプレイが行われているサービスの一つが前述の人狼 **BBS**（図 1）[人狼 BBS] である。人狼 **BBS** ではこれまでに 2 000 回以上のゲームが行われており、大量のログが蓄積されている。

人狼知能プロジェクトでは、**BBS** タイプの人狼を対

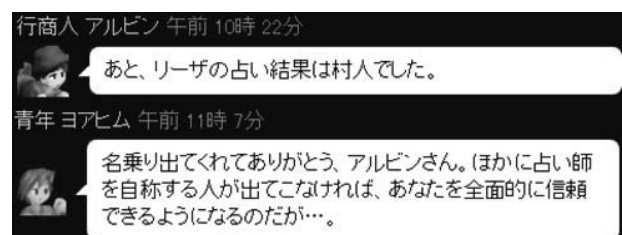


図 1 人狼 BBS における占い師 CO のシーン [人狼 BBS]

象に考え、基本ルールを人狼 BBS に準拠させることとする。それに伴い、人狼 BBS におけるプレイログをエージェントの設計などに利用することが可能となる。

2.6 対面人狼

日本においては、一般的な人狼ゲームとして後述するタブラの狼[タブラの狼]やうそつき人狼[うそつき人狼]など多数のカードゲームが販売されている。これらのカードなどを用いて、10人程度の人がテーブルを囲んで行うのが対面人狼と呼ばれる一般的な人狼のスタイルである。種類によって、さまざまな特殊な役職やルールが用意されており、楽しめるような工夫がされている。中には、究極の人狼[究極の人狼]のように、最大74人で楽しめるものや、人狼系なりきり推理ゲームダンガンロンパ1・2超高校級の人狼[ダンガンロンパ]などゲームやアニメとコラボレーションされたものも用意されている。

対面人狼では、論理的な議論ももちろんのこと、特に理由はよくわからないが、なんか怪しいといった、仕草や身振り手振りなど、ノンバーバルな情報も重要視される。例えば、自分が人狼で3人が残り、論理的に破綻しかけていても、全身全霊で否定し、別の可能性を匂わせることで、自分を人狼だと強く思っていた人が、土壇場で判断を変えることもよくあることであり、そこが対面人狼の面白いところでもある。BBS人狼は、ロジックがすべてのところがあり、論理的な思考と説明ができるプレーヤーはおしなべて強いプレーヤーとなる。しかしながら、BBS人狼で強いプレーヤーでも対面でやると、しばしば論理的に正しくても勝つことができない。それを単に運が悪かったと感じるプレーヤーもいるが、感情面やノンバーバル情報を利用した駆け引きを楽しむのも対面人狼の大きな特徴であるといえる。

2.7 エンターテインメントとしての魅せる人狼

オンライン型人狼や、対面人狼などの一般的な議論のスタイルで行う人狼に限らず、さまざまなスタイルの人狼がこれまでに登場している。本章では、エンターテインメントとしての魅せる人狼について説明を行う。

§1 人狼の演劇

人狼を用いた演劇として有名なものに人狼 TLPT [人狼 TLPT] の存在がある。人狼 TLPT とは、人狼ザ・ライブ・プレイング・シアター (The Live Playing Theater) の略で、ステージ上で13人の役者がアドリブ演技を繰り返しながら、人狼をプレイしそれをエンターテインメントとして披露する人狼スペシャリスト集団のことである。役者達が村人を演じながら人狼をプレイし、毎回筋書きのないドラマをアドリブでつくり上げていく舞台であり、まさに「魅せる人狼」というジャンルをつくり上げた存在である。役者が演技を行う舞台ではなく、また、単に役者が人狼を行うのでもない、全く新しい形のエン

ターテインメントになっている。普通の舞台と違うのは、すべてアドリブで毎回ストーリーもエンディングも変わるため、公演内容のネタバレは許可されており、公演中も Twitter などでも実況してもよいというルールがある。また観客も一緒に見ながら誰が人狼かを予想することができる。観客も一体となって狼探しをすることで、客席も村の一員という雰囲気をつくっている。

人狼 TLPT は、2012年から始まっており、出演者は現時点で176名で、これまでに数多くの公演が行われている。一部の公演はニコニコ動画で有料で視聴が可能になっているので、ぜひ一度見てほしい。また、役者による舞台だけではなく、将棋棋士との対戦や、ゲームクリエイターとの対戦なども行っており、その様子も公開されている。

§2 人狼のテレビ番組と動画

2013年は人狼のブームになったが、その先駆けの存在として、テレビ番組『人狼～嘘つきは誰だ?～』(フジテレビ系)[人狼～嘘つきは誰だ?～]の影響が大きい。フジテレビの人狼では、芸能人が、毎回アドリブでしのぎを削り、面白い対戦をつくり上げていく。番組のセットや、ルール設定、素晴らしい編集のおかげもあるが、芸能人の対話や反応も大きな要素である。こちらも魅せる人狼を意識した、興味深い試みの一つであるといえる。

3. 人狼知能プロジェクトと AI

著者らは、人狼の人工知能分野における将来性に注目し、共同研究プロジェクトである「人狼知能プロジェクト」を立ち上げた。本章では、人狼知能プロジェクトとプロジェクトが目指す目標について詳しく説明する。

3.1 人狼知能プロジェクトとは

人狼知能プロジェクトは、2013年に人工知能・自然言語処理・HAI 研究者らによって立ち上げられた、人狼をプレイする人工知能を実現するためのプロジェクトである。人狼知能プロジェクトでは、人狼知能プロトコルと人狼知能サーバの情報を公開することで、多くの研究者や開発者がエージェントの作成に参加できる環境を整える。さらに、これらのエージェントを集めたゲーム大会を開催する。これによって、人狼知能を汎用人工知能の新しい標準問題とすることを目指している [篠田 14]。

3.2 人狼ゲームの性質

人狼ゲームを標準問題として取り扱う理由として、まず人狼ゲームの性質 [大澤 13] について説明する。

§1 客観的視点での情報非決定性

人間同士で行うゲーム、特にボードゲームは、いくつかの種類に分かれている。その中で「コミュニケーションゲーム」と総称される分野がある。コミュニケーションゲームではゲームの勝敗はプレーヤー同士の情報交換に

依存して決定される。人狼は、このようなゲームの中で最も極端な形をもつゲームであり、コミュニケーション以外の客観的情報、勝敗決定要因がほとんど存在しない、という性質をもっている。人狼で客観的といえる情報は、各プレイヤーの発言内容自体と、日数、処刑者、襲撃者などであり、人狼に登場するほとんどの役職では、自分自身の役職を「客観的に」証明する手段をもたない。

具体例をあげてみよう。例えば、あなたが「村人」であり、あなたの友人がそのゲームを外から観察している（ゲーム上の昼の発言のみを観測している）とする。このゲームプレイ中に、あなたが村人、あるいは人間側であることをその友人に対し証明することは不可能である。これは、あなたがどの役職であっても同様である。もちろん、ゲーム中の「人狼」のプレイヤーにとってあなたが少なくとも「非人狼、つまり人間側」であることは自明であるし、もし村の中の占い師があなたを占ったとすると、その占い師はあなたが人間側であることを知ることができ、また占い師はあなたが人間であるという発言を行うことができる。しかしながら、村の外にいるあなたの友人から見たとき、「人狼」のプレイヤーが人狼であるのか、あるいは「占い師」と宣言した人間が本当に占い師であるのか、そしてあなたが本当に人間であるのかどうかはゲーム終了までわからない（図2）。

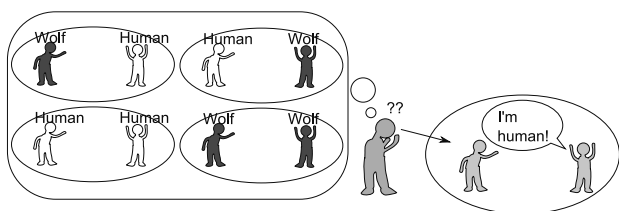


図2 客観的視点での情報非決定性の例。
あるプレイヤーに人狼だと疑われたときにあなたが人間側であると発言した場合において、そのゲームを外から観察している友人の視点からは複数の可能性が考えられ、あなたの発言が正しいことを証明することはできない

人狼というゲームにおいて情報の共有は個人、もしくは数人の間にとどまっており、参加者全員で共有できる情報はない。ほとんどの人狼のバリエーションにおいて、客観的視点から情報が決定できない、という原則は守られている。これは人狼というゲームの基本原則であるといっていよい。また、プレイ中に客観的視点でいっさい情報が決定しないゲームは、人狼とそのバリエーション以外では滅多に存在しない。この客観的視点での情報非決定性は、数あるコミュニケーションゲームの中でも、人狼に特有の性質と考えてよい。

§2 プレーヤの非思考委任性

プレーヤは、客観的視点での情報非決定性により、自身の証明を自身の言語や態度で行い、自身の決断を自身のみの判断で行う、という状況に置かれる。そのため、プレーヤは誰かを盲信し、自分の選択をその人に委ねる、という手段を取ることが難しい。我々が日常生活を営む

うえで自然と行っている思考の委任（例えば『先生が言っていたから、偉い人が言っていたから、専門家が言っていたから』という思考）が許されないし、客観的な知識ベースに頼ることもできないということを意味する。人狼において行われる会話では、客観的に役職を特定できる会話はなく、すべては文脈に依存しているといえる。例えば、村人として正しい発言をしているからといって、そのプレーヤが人間であるということにはならない。実際のプレイにおいては、〇〇という発言は村人らしい、というようなセオリーと呼ばれるヒューリスティクスは存在するが、これはあくまでコミュニティに依存した情報であり、違う人狼コミュニティではセオリー自体が異なることもある。

プレーヤは自身の役職を客観的には証明できないし、他のプレーヤの役職が何であるかを客観的に決めることも原理的にはできない。では人狼において、プレーヤはどのように自らの意思を決定し、選択を行うべきなのだろうか。

§3 推理：他者の意図のモデル化

発言の内容を支える客観的な情報が存在しない場合、話者の背景が貴重な情報源となる。つまり「相手の立場に立つ」ということである。人狼で交換される情報は客観的な情報ではなく、どこまでも主観的な情報（すなわち、文脈に深く依存した情報）のみである。人狼ゲームにおけるすべての重み付けは、最終的にプレーヤ自身が行わなければならない、それはプレーヤ自身にしか決定できない（図3）。

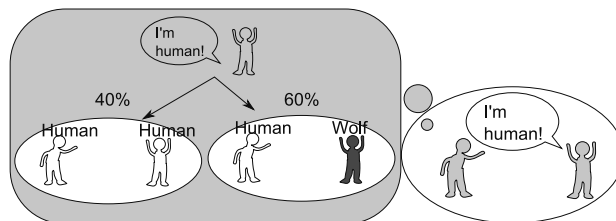


図3 プレーヤによる推理の例。
プレーヤはそれまでの文脈と、発言者の発言と立場を考慮したうえで、推理を行い、発言の正しさの尤度を決定する

占いや霊能結果は、相手の発言意図を推し量る貴重な手掛かりとなる。もしそれが偽物であったとしても、偽物の結果は偽者の手掛かりを与えてくれる。もちろん本物の占い師が本当のことを言うとも限らないし、本物の占い師にとって嘘を付くことが必要な局面も存在する。

村人は情報を集め、人狼は情報を誤魔化す。村人は可能性の枝を刈り、人狼は可能性の枝を育てる。大まかにいえば、人狼は人間と人狼の間の情報の争奪戦であり、論理で枝を示し、説得で枝を刈る、あるいは増やす。

ただし、人狼が最善手を打った場合、可能性の枝は、結局一つにはならない。人狼は誰が人間かを知っている。つまり、人狼は占い師や霊媒師を『正確に真似る』ことができる。いかに情報を積み重ねたとしても、この原則

は変わらない。本物と偽物を見分ける術は、場に出た情報・言語のみでは与えられない。

§4 説得：他者から見た自己のモデル化

情報の確かさが立場により異なるため、最終的な尤度はプレーヤにより異なり、結論も個々で異なる。しかし人狼では日ごとに、村全体の決断として最終的に処刑先を決定しなければならない。決定先は個々人が独立に判断して投票する自由投票も存在するが、たいていはその前に意見を集約するための議論を行う。

互いに異なる意見をもつプレーヤ達は、最終的に自分達の意見を集約し、合理的な決断を下す必要がある。ここでは推理よりも、説得や信頼といった要素が重要なキーワードとして浮かび上がってくる。人間も、人狼も、生き残りたければ信頼を勝ち取らねばならない。自分の意見を通すためには、自分の意見が信頼に足るものであることを他者に説明する必要がある。この段階では、他者のモデル化だけではなく、他者から見た自己のモデル化、信頼できる発言の演出が重要となってくる(図4)。

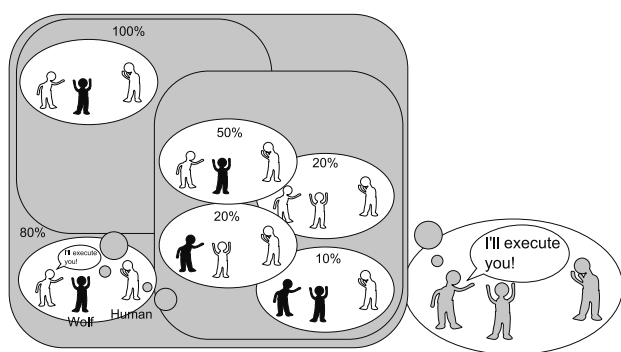


図4 他者から見た自己のモデル化の例。
自分があるプレーヤを人狼だと思った場合、自分以外のプレーヤが、自分をどうモデル化しているかも検討したうえで発言する必要がある

他者説得のために行われる言語合戦はさまざまな手法があり、最後は、これまでに得た情報や発言による論理的な説得ではなく、感情に訴えることで、最終局面において相手の判断を変え得ることも多い。

3.3 汎用人工知能の標準問題としての人狼

標準問題とは、ある技術の性能などを評価するうえで、その善し悪しを比較可能とするために設ける問題のことである。人工知能が対象とする、現実存在する諸問題は非常に幅広く多様である。そのため、同じ領域を対象として課題を設けたとしても、必ずしも同様の環境を構築できるとは限らず再現性に乏しい。そこで、互いがそれぞれの問題を直接調べる代わりに、現実におけるさまざまな課題の中で、重要かつ一般的な問題を十分に表現でき、現実の問題へのアプローチとなる問題を標準問題としてきた。

これまで、人工知能の標準問題としてさまざまな問題が提案されてきている。代表的なものとしては、チェス

や将棋・囲碁、追跡問題、囚人のジレンマ、RoboCupなどがある。チェス・将棋・囲碁などはゲーム進行のスコアリングと探索の問題であり、追跡問題は共通の全体ゴールをもつ集団が全体の戦略と個々の戦術との関係を追求する問題である。そして、RoboCupでは、ソフトウェアとしての人工知能技術とロボットというハードウェアを融合しサッカー・災害救助・日常環境という実社会への投入を前提として問題を提案している。これらに共通するのは、それぞれの課題に対して、共有すべき大きな目標を設定し、共通のフレームワーク上で互いの技術・知見を実装することで、利点欠点を比較可能となる。

汎用人工知能(Artificial General Intelligence, 以下 AGI)とは、人間レベルの知能を実現することを目指すとした人工知能研究のことである。この汎用人工知能の具体的な姿は人工知能や人間の汎用知能と同じく明確な定義は存在しないが、Adamsらにより、汎用人工知能の実現に向けた課題や評価シナリオが整理されている[Adams 11]。汎用人工知能に関する詳細な説明は、本稿では割愛するが、Adamsらは、汎用人工知能の評価のためのシナリオとして六つのシナリオを提案し、これらのシナリオをベースとして、必要とする環境とそこで行うべきタスクを設定することを求め、AGIの実現に必要とされる能力の領域を提案している。具体的には、知覚、記憶、注意、社会インタラクション、プランニング、モチベーション、推論、運動作業、コミュニケーション、学習、情動、自己・他者モデル、構築・創造、定量化など15の能力エリアを設けている。

人狼知能プロジェクトでは、この人狼ゲームを、ソフトウェアエージェントを用いたゲームとして、将来的には、人間を交えたゲームとして準備をしている。

人狼ゲームにおいて、プレーヤをソフトウェアエージェントとして対戦する人工知能という観点から見たときに、その人工知能(人狼知能)を実現するには、その研究分野は、エージェントやインタラクション・認知科学・心理学、ロボット工学など多様な領域にまたがる。Adamsらのあげた能力エリアでいうと、ソフトウェアエージェントとして実装するだけでも、記憶、社会的インタラクション、プランニング、推論、コミュニケーション、学習、自己・他者モデル、定量化など多岐にわたる能力が必要となる[Adams 11]。

これに、人間を交えた実世界でのゲームを行うとした場合には、知覚器官や運動能力なども加わる。視覚情報、自然言語処理、ロボットの実装などの基本的なところはもちろん、高度な会話能力や表情、社会性など、人間として高度な能力エリアの大半をカバーする課題といえる。

ここで、標準問題としての面白さを考えたときに、知能の発展に比例して、ゲームの面白さが破壊される場合がある。例としては、将棋など手数を読み切れる終盤における完全情報ゲームなどがあげられている[伊藤 14]。

しかし、人狼のような複数のプレーヤが参加するコミュニケーションゲームでは、複数の新規プレーヤによる他者モデルの見積もりは完全に読み切ることは困難であるため、知能が高度化することによって、ゲームの面白さが破壊される影響は少ない。これは、熟練プレーヤの中に初心者プレーヤが数名混じることにより、熟練プレーヤの思惑どおりにはゲームが進まないことが良い例となるだろう。

3.4 これまでに得られた面白い結果と5人人狼のすすめ

人狼知能プロジェクトでは、各分野の研究者が、それぞれの専門を生かし、これまでにBBS人狼のログデータを用いた分析[稲葉 12, 稲葉 13]やデータからの学習の研究[梶原 14, 鳥海 14b], 人狼プロトコルの開発[大澤 13, 大澤 14], 対面人狼における仕草やノンバーバル情報の研究[片上 14, 高久 13]を行ってきている。現在では人狼サーバを公開し、人狼知能エージェント同士や人間と人狼知能エージェントとの対戦が可能になっている。

また、人狼知能プロジェクトに関連する研究において、これまでに得られた面白い結果を以下に載せておく。ぜひ実際のプレイにあたって参考にしてほしい。

- プレーヤの残り数が5人もしくは7人のとき、人狼プレーヤは誰も襲撃しないほうがよい(人狼AI)[鳥海 14a]
- 雑談すると殺される(人狼BBS)[稲葉 13]
- うなずく、顔を触る、腕を組むと殺される(対面人狼)[片上 14]

人狼知能プロジェクトでは、将来的に人間と対面プレイが可能な物理エージェント(ロボットや擬人化エージェント)を構築することが本プロジェクトの目的の一つとなっており、人間の仕草の分析の研究や、人間と同様の仕草を実装した、擬人化エージェントによる人狼対戦システムが開発されている(図5)[片上 14, Kobayashi 14]。

また、これまでに、片上らは研究対象として「5人人狼」(いわば詰め将棋の人狼版)の設定を開発・採用した[片上 14, Kobayashi 14]。プレーヤは5名で行う、人狼の魅力を凝縮した少人数人狼であり、本稿では、これを5人人狼と呼ぶことにする。役職は、人狼、狂人、占い師、村人2名であり、初日の夜に占い師が一人を占ったところからスタートする。ゲームとしては、昼に3～5分の議論を行い、1日もしくは2日で終わるゲームであるので、10～20分ほどで終わってしまうが、この簡単な設定においてもさまざまな可能性が考えられ、また、どの可能性も五分五分の状況が多く、また一度の失敗も許されないため、村人も含めて5名がフルに頭を悩ませる、とてもエンターテインメント性の高い人狼設定である。人狼プロジェクトでは、この5人人狼において、真の占い師が、実際には占っていない村人Aに対して、「Aは

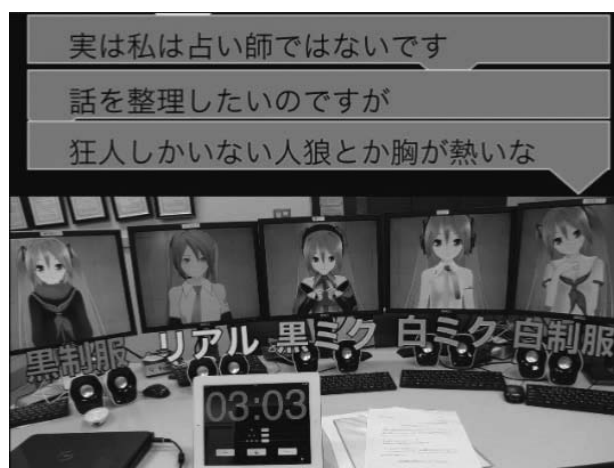


図5 擬人化エージェントによる5人人狼の様子[片上 14]。モデルは初音ミク5人を利用し、MMDAgentを用いて実装した人狼対戦システムにおいて、Woz方式にて、対戦を行っている。場面は、リアルミクが占い師COを行い黒ミクを人狼判定し、白ミクも占い師COをして白制服ミクを人狼判定したが、途中で白ミクが占い師COを取り消したケース。この後の展開は、今後動画を人狼知能Webサイトにて公開する予定

占った結果人狼でした」と言うことによって、村人側が勝利に近づくという、通常の人狼では、あり得ない戦略が生まれることを発見し、人数や設定によって戦略性が大きく異なることを説明している。5人人狼は気軽にできるゲームなので、なぜそうなるのかについて考えてみてほしい。

3.5 人狼による教育

人狼は人を騙すゲームだと揶揄していわれることがまれにあるが、実は説得のためのゲームでもある。自分の役になりきり、自分の主張を論理的に説明することで、いかに相手に納得してもらうかを競う言葉のスポーツでもある。自分の本来の主張とは異なる可能性もある役割に割り振られて、その立場に立って議論を行う点でディベート教育と近いが、ディベートにおいては、事前の調査が結果に大きく影響する。人狼においては調査の必要はなく議論で勝ち負けを決めることにより、純粋に議論の善し悪しの評価を手軽に行える点で大きく異なる。また、人狼の特性の一つでもある、非思考委任性により、自身の考えを他人の考えに委ねることなく、自分で考え、自分で判断することが鍛えられる。

大学教育において、例えば、工学部の授業においては、学生ははじめに授業に出て課題やテストなど必要なものを提出していれば、それなりの成績で単位を取れることも多いが、その一方で、就職活動では、グループ面接や役員面接などさまざまな社会的な能力を必要とし、それまで成績の良かった学生が、うまく成果を出せず、苦戦することも多々ある。著者の片上は、東京工芸大学工学部コンピュータ応用学科においてPBL型授業の一環として、上記の特性をもつ人狼を用いた対話教育を行っている。人狼による対話教育も、人狼知能プロジェクトの

視野に入っている。

3・6 人狼知能プロジェクトにおけるイベントなど

広報活動と今後の予定

人狼知能プロジェクトにおける最近の広報活動と今後の予定に関して、以下に説明する。

CEDEC 2014 [CEADEC 14] において「将棋の次は人狼か?」と題して、パネルディスカッションが行われた。人狼知能プロジェクトからは、松原、大澤、稲葉が参加し、さらに人狼の魅力にはまっているゲームクリエイターであるイシイジロウ氏が人狼と人工知能の可能性について語った。このパネルセッションにおいては、ファミ通 [ファミ通] や 4Gamers [4Gamers] でも詳しく解説されており、人狼について興味がある方は、ぜひ一読されたい。

一般の雑誌 [新潮 14] において、人工知能が人間を超える「2045年」問題として、人狼プロジェクトの一人である松原と科学作家の竹内 薫氏の対談が掲載されており、科学研究の次なる標的は人狼であり、15年か20年かはかるといわれるが、それができたら胸を張ってチューリングテストを通ったと言ってよいと語っている。

エージェントの合同シンポジウム JAWS 2013 [JAWS 13] および JAWS 2014 [JAWS 14] では、イベントとして、人狼知能プロジェクトによって、人狼大会の運営と人狼クイズ大会運営を行ってきた。今年度は、人狼知能 AI 同士の対戦を予定している。

また、2014年11月19日に本学会の公募型 AI セミナーにて、「人狼知能研究のすすめ」[JSAI 14] と題して、人狼知能エージェントの研究環境の構築方法を、人狼プロジェクトの講師陣が解説、実演する予定である。

4. お わ り に

本稿では、人狼をプレイする人工知能を実現しようと試みる人狼知能プロジェクトの紹介を行い、エンターテイメントとしての人狼、AI 領域における人狼、そして教育における人狼などの視点を通して、人狼知能プロジェクトが目指す人狼の魅力について説明してきた。本稿によって、人狼知能について興味をもち、人狼をプレイする人、そして AI プレーヤーを構築してくれる人が一人でも増えたら、嬉しい限りである。人狼プロジェクトにご興味をもたれた方は、プロジェクト HP (<http://www.aiwolf.org/>) をご覧いただければ幸いです。

◇ 参 考 文 献 ◇

- [Adams 11] Adams, S. S., et al.: *Mapping the Landscape of Human-Level Artificial General Intelligence*, AAAI (2011)
[aduma 03] 「汝は人狼なりや?」と小一時間問い詰めた (2003). <http://web.archive.org/web/20031123075546/>

- <http://park1.wakwak.com/~aal/okami/>
[CEADEC 14] *Computer Entertainment Developers Conference (CEDEC 2014)* (2014). <http://cedec.cesa.or.jp/2014/>
[ダンガンロンパ] 株式会社アルジャーノンプロダクト: 人狼系なりきり推理ゲーム ダンガンロンパ 1・2 超高校級の人狼, http://www.algernonproduct.co.jp/werewolf/drl_2.htm
[ファミ通] 人間をだます「人狼知能」の実現は5年後、いや15年後、もしや50年後!? 難題に挑む人工知能研究者とゲームクリエイター・イシイジロウ氏が激論!, *CEDEC 2014* (2014). <http://www.famitsu.com/news/201409/04060581.html>
[稲葉 12] 稲葉通将, 鳥海不二夫, 高橋健一: 人狼ゲームデータの統計的分析, *Game Programing Workshop*, No. 6, pp. 144-147 (2012)
[稲葉 13] 稲葉通将, 大島菜央実, 鳥海不二夫, 高橋健一: 雑談ばかりしていると殺される一人狼 BBS におけるプレイヤの発言傾向と意思決定・勝敗の分析一, *JAWS 2013* (2013)
[伊藤 14] 伊藤毅志, 杵渕哲彦, 藤井叙人: ゲームにおけるヒューマンエラー—将棋による考察—, *19th Game Programming Workshop*, pp. 196-201 (2014)
[JAWS 13] 合同エージェントワークショップ&シンポジウム 2013 (2013). <http://jaws-web.org/event/jaws2013/>
[JAWS 14] 合同エージェントワークショップ&シンポジウム 2014 (2014). <http://jaws-web.org/event/jaws2014/>
[JSAI 14] 2014年人工知能学会・公募型 AI セミナー「人狼知能研究のすすめ」(2014). <http://www.tkomat-lab.com/kikaku/>
[片上 14a] 片上大輔, 金澤将志, 小林 優, 大澤博隆, 鳥海不二夫, 稲葉通将, 篠田孝祐: 擬人化エージェントを用いた人狼対戦システムのための仕草分析, *JAWS 2014* (2014)
[片上 14b] 片上大輔, 小林優, 大澤博隆, 稲葉通将, 篠田孝祐, 鳥海不二夫: 擬人化エージェントを用いた人狼対戦システムの開発, *Game Programing Workshop*, P-9 (2014)
[梶原 14] 梶原健吾, 鳥海不二夫, 大橋弘忠, 大澤博隆, 片上大輔, 稲葉通将, 篠田孝祐, 西野順二: 強化学習を用いた人狼における最適戦略の抽出, 情処第 76 回全国大会 (2014)
[Kobayashi 14] Kobayashi, Y., Osawa, H., Inaba, M., Shinoda, K., Toriumi, F. and Katagami, D.: Development of werewolf match system for human players mediated with lifelike agents, *2nd Int. Conf. on Human-Agent Interaction (HAI2014)* (2014)
[究極の人狼] アークライトゲームズ: 究極の人狼第2版, 完全日本語版. <http://www.arclight.co.jp/ag/agbg/agbg.php?code=LG-0015&type=b>
[汝は人狼なりや? 01] 汝は人狼なりや?, Loony Labs. (カードゲーム) (2001)
[大澤 13] 大澤博隆: コミュニケーションゲーム「人狼」におけるエージェント同士の会話プロトコルのモデル化, *HAI シンポジウム 2013*, pp. 122-130 (2013)
[大澤 14] 大澤博隆, 鳥海不二夫, 片上大輔, 篠田孝祐, 稲葉通将: 人狼ゲームのプロトコル設計: 推理と説得のプロトコル, *FAN2014* (2014)
[篠田 14] 篠田孝祐, 鳥海不二夫, 片上大輔, 大澤博隆, 稲葉通将: 汎用人工知能の標準問題としての人狼ゲーム, 人工知能学会全国大会, 2C4-OS-22a-3 (2014)
[新潮 14] 竹内 薫, 松原 仁: 人工知能が人間を超える / 2045 年問題, 週刊新潮, Vol. 59, No. 39, pp. 134-136 (2014)
[タブラの狼 02] タブラの狼, daVinci (カードゲーム) (2002)
[高久 13] 高久奨乃, 片上大輔: 人狼ゲームにおいてノンバーバル情報がプレイヤーに与える影響について, *JAWS2013* (2013)
[鳥海 14a] 鳥海不二夫, 梶原健吾, 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐, 西野順二: 人工知能は人狼の夢をみるか—人狼知能プロジェクト, デジタルゲーム学会 2014 (2014)
[鳥海 14b] 鳥海不二夫, 稲葉通将, 大澤博隆, 片上大輔, 篠田孝祐: 人狼における学習, *FAN 2014* (2014)
[うそつき人狼 13] うそつき人狼, 株式会社人狼 (カードゲーム) (2013)
[4Gamers] 将棋でプロを倒した AI が、次に狙うのは人狼。「人狼知能」プロジェクトに迫る, *CEDEC 2014* (2014). <http://www.4gamer.net/games/999/G999905/20140905115/>
[人狼 BBS] 人狼 BBS (ブラウザゲーム). <http://ninjinix.com/>
[人狼知能プロジェクト 13] 人狼知能プロジェクト Artificial Intelligence

based Werewolf (2013). <http://www.aiwolf.org/>
 [人狼 Online 14] 人狼 Online(2014). <http://jinrou-online.com/>
 [人狼 TLPT] 人狼ザ・ライブプレイングシアター. <http://7th-castle.com/jinrou/>
 [人狼読本 13] 人狼読本, KADOKAWA enterbrain (2013)
 [人狼～嘘つきは誰だ?～] 人狼～嘘つきは誰だ?～, フジテレビ.
<http://www.fujitv.co.jp/jinroh/>

2014 年 11 月 14 日 受理

著 者 紹 介



片上 大輔 (正会員)

2002 年東京工業大学大学院総合理工学研究科知能システム科学専攻博士後期課程修了, 博士 (工学). 同年, 東京工業大学大学院総合理工学研究科助手. 2007 年同研究科助教. 2006 年英国ハートフォードシャー大学, スイスチューリヒ大学の客員研究員を兼任. 2010 年東京工芸大学工学部コンピュータ応用学科准教授. 現在に至る, 日本知能情報ファジィ学会, IEEE, ACM 各会員. HAI に関する研究に従事し, 人狼知能および雰囲気工学に興味をもつ.



鳥海 不二夫 (正会員)

2004 年東京工業大学大学院理工学研究科機械制御システム工学専攻博士課程修了. 同年, 名古屋大学情報科学研究科助手, 2007 年同助教, 2012 年東京工業大学大学院工学系研究科准教授, 現在に至る. エージェントベースシミュレーション, 人工市場, ソーシャルメディア, 人狼知能などの研究に従事. 電子情報通信学会, 日本社会情報学会各会員, 博士 (工学).

大澤 博隆 (正会員) は, 前掲 (Vol. 30, No. 1, p. 6) 参照.



稲葉 通将 (正会員)

2012 年名古屋大学大学院情報科学研究科博士後期課程修了, 同年, 広島市立大学大学院情報科学研究科助教. 現在に至る. 対話システム, 対話処理, 対話ゲームに関する研究に従事. 博士 (情報科学). 電子情報通信学会, 情報処理学会, IEEE 各会員.



篠田 孝祐 (正会員)

1999 年名古屋工業大学知能情報システム学科卒業. 2004 年北陸先端科学技術大学院大学知識科学研究科博士後期課程修了. 現在, 電気通信大学大学院情報システム学研究科助教. 理化学研究所 客員研究員, 文部科学省科学技術・学術政策研究所客員研究官を兼任. 博士 (知識科学). 複雑ネットワーク, 社会シミュレーションなどの研究に従事. 情報処理学会

会員.



松原 仁 (正会員)

1981 年東京大学理学部情報科学科卒業. 1986 年同大学院工学系研究科情報工学専攻博士課程修了. 同年, 通産省工業技術院電子技術総合研究所 (現 産業技術総合研究所) 入所. 2000 年公立はこだて未来大学教授. ゲーム情報学, エンターテインメントコンピューティング, 観光情報学, 公共交通システムなどに興味をもつ. 著書に「将棋とコンピュータ」(共立出版, 1994), 「ロボット情報学ハンドブック」(近代科学社, 2010), 「コンピュータ将棋の進歩」(共立出版) など. 本学会会長, 情報処理学会理事.