

人狼知能プロジェクト **特** 集 「エンターテイメントにおける 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/ 東京大学大学院工学系研究科

鳥海 不二夫 Fujio Toriumi 大澤 博隆

Hirotaka Osawa

Department of Applied Computer Science, Faculty of Engineering, The University of Tokyo. tori@sys.t.u-tokyo.ac.jp 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 広島市立大学大学院情報科学研究科 Graduate School of Information Sciences, Hiroshima City University Michimasa Inaba inaba@hiroshima-cu.ac.jp 篠田 孝祐 電気通信大学大学院情報システム学研究科 Kosuke Shinoda Graduate School of Information Systems, The University of Electro-Communications

松原 公立はこだて未来大学システム情報科学部 Hitoshi Matsubara Faculty of Systems Information Science, Future University Hakodate. Keywords: werewolf AI, werewolf, communication game, general artificial intelligence.

ロジェクト [人狼知能プロジェクト 13] を立ち上げた. 1. は じ め に

本稿では、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、さまざ 「ここは、とある村、平和に暮らしていた村人達にあ まな視点を通して、この人狼知能プロジェクトが目指す る日, 衝撃的な事実が告げられる. この村の住人の中に, 人狼について説明する.

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

66

2.2 人狼ゲームの流れ

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、人狼をすべて処刑することができれば、村人達は 起こし始めている。著者らは、この人狼のもつ魅力に注 する. 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

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

65

平和な村を取り戻して勝利. 生き残っている人狼の数 界中でプレイされている. 日本においてもタブラの狼

が村人と同数になれば、村は滅んで人狼の勝利となる.」 [タブラの狼 02] やうそつき人狼 [うそつき人狼 13] など 以上が、コミュニケーションゲーム「人狼」の大まかな 多数のゲームが販売されている. 人狼をプレイする方 世界観である[人狼読本13]. 古くからヨーロッパでエ 法としては、前述したような市販のカードなどを使って ンターテイメントとして行われてきたこのゲームが、こ テーブルを囲んで行う対面型と、電子掲示板などの Web こ数年, さまざまなメディアを通して、日本でブームを 上のアプリケーションを使って行う BBS タイプが存在

す.人間陣営の目標は人狼の全滅に、人狼陣営の目標は うかを知ることができる。 人間の人数を人狼の人数と同数以下にすることにあり, (4) 狩人 人間陣営に所属する. 夜のフェーズでプレーヤー人を 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 指定して、人狼の襲撃から守ることができる。狩人が守 以外の誰がどの役職かわからない。特に、人間側は誰が ろうとしたブレーヤを人狼が襲撃した場合、その日は誰

る. ブレーヤは役職によって、人間または人狼陣営にそ (3) 霊媒師(霊能者)

ブレーヤにはまずランダムに「役職」が割り当てられ る. 人間陣営においては最も重要な役職である.

ヤは同じ人狼陣営のプレーヤ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は互いに協 コミュニケーションを取ることが可能である(対面の人 力しながら、人間陣営に正体がばれないように行動する 狼の場合、夜のフェーズで目配せとジェスチャによるコ 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 を襲撃することが可能. フェーズではすべてのプレーヤによって、誰が人狼かを (6) 狂人 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は当該能力に あるかは把握されず, 狂人からも誰が人狼なのかはわか よって知り得た情報を用いて、自分達の陣営が有利にな らない、それ以外は特殊能力がないただの村人と同様で るように議論を導くことになる. 一定期間の議論の後, あるが, 人狼陣営が勝利したときに勝利となるため, ひ プレーヤ全員の投票によって、人狼と考えられる人物を そかに人狼に協力をする、勝利人数のカウント時には人

処刑する.処刑されたプレーヤはゲームから除外され,

夜のフェーズでは、人狼陣営に所属するプレーヤが人

ゲーム終了まで参加することができない.

イとなる.

れぞれ振り分けられ、各プレーヤはチームの勝利を目指

とが基本的な行動指針となる.一方、人狼陣営のプレー (5)人狼

間陣営のプレーヤを一人選び、襲撃する、襲撃されたプ 2.4 人狼ゲームの基本戦略 レーヤは死亡者として扱われ、処刑されたプレーヤと同 §1 人間陣営の基本戦略 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. る.人狼を探し出すには占い師と霊媒師による情報が重 昼のフェーズと夜のフェーズを繰り返し、勝利陣営を決 要である.しかしながら、多くの場合人狼陣営のプレー 定する. 議論において、人間陣営に所属するプレーヤは人狼の CO: Coming Out) するため、誰が真の占い師、霊媒師 嘘を見破るかが最大のポイントとなる.また.能力をも であるかを見抜くことが必要となる. つ役職に就いたプレーヤは能力によって知り得た情報を 使って他のプレーヤを説得することがポイントとなる. が自分を人狼であると指摘した場合, その占い師は人狼 一方、人狼陣営のブレーヤは自分達が不利にならないよ 陣営(人狼もしくは狂人)に所属していることがわかる. うに議論を誘導し、時には能力をもった役職であると偽

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 なくなるようにプレイすることが多い。 ルや、バージョンが存在し、役職もそれによって大きく §2 人間陣営役職もちの基本戦略 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

人狼知能プロジェクト 占い師であることは隠しておくことが多い。ただし、狩

とが必要となる.

する役割になることも多い.

することも可能である.

§3 人狼陣営の基本戦略

人が残っていることがわかっている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 プレーヤ(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ーヤを説得するこ

霊媒師は、処刑されたプレーヤ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、人狼陣営のプレーヤが霊 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を

狩人は他のプレーヤ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、

人狼陣営の基本戦略は人間陣営を騙して人狼陣営の

プレーヤが処刑されないように議論を誘導することにあ る.人狼陣営のプレーヤ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ーヤが信じないよう議論を誘導す

人狼かわからないため、会話の中から人狼を探し出すこ も死なないことになる. 人狼陣営に所属する.複数人狼がいる場合は,互いに

人 工 知 能 30巻1号(2015年1月)

指定した一人が「人狼であるか否か」を知ることができ

人間陣営に所属する.前日に処刑した人物が<mark>人狼かど</mark>

ミュニケーションを行う). 夜のフェーズで任意の村人

人狼陣営に所属する。ただし、人狼からは誰が狂人で

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. 人間陣営の基本戦略は人狼陣営の嘘を見抜くことにあ ヤが偽の占い師および霊媒師として名乗り出る(以下

例えば, 占い師 CO を行った二人の占い師のうち一人

このような情報を積み重ねていくことによって、各ブ り、議論を間違った方向へ誘導することなどが基本プレレーヤ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 狼陣営のプレーヤが人間陣営のプレーヤを騙す要素が少 人間陣営で役職をもつプレーヤは占い師, 霊媒師, 狩

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.

67 2. オンライン型人狼(短期人狼, BBS 人狼) 一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

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」でのプレーヤ同士の議論と、人狼が処刑を決

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人のプレーヤが人狼を 行う ROOM に入り、プレーヤ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

める「夜」をチャット上で再現し人狼を行う.

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、守られたプレーヤが 人間陣営であることがわかるため、他のプレーヤよりも 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

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ーヤのう 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ー ヤ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 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.

68 象に考え, 基本ルールを人狼 BBS に準拠させることと ターテイメントになっている. 普通の舞台と違うのは, する. それに伴い、人狼 BBS におけるプレイログをエー すべてアドリブで毎回ストーリもエンディングも変わる

ジェントの設計などに利用することが可能となる.

日本においては、一般的な人狼ゲームとして後述する

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

が用意されており、楽しめるような工夫がされている.

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

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

に運が悪かったと感じるプレーヤもいるが、感情面やノ

タブラの狼[タブラの狼]やうそつき人狼[うそつき人狼] 村の一員という雰囲気をつくっている.

2.6 対 面 人 狼

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大にある. 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2000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 あと、リーザの占い結果は村人でした。 **青年 ヨアヒム** 午前 11時 7分 名乗り出てくれてありがとう、アルビンさん。ほかに占い師 を自称する人が出てこなければ、あなたを全面的に信頼 できるようになるのだが…。 図1 人狼 BBS における占い師 CO のシーン [人狼 BBS]

人 工 知 能 30巻1号(2015年1月) 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

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 である.種類によって、さまざまな特殊な役職やルール 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 中には、究極の人狼[究極の人狼]のように、最大74人 エータ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

著者らは、人狼の人工知能分野における将来性に注目

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. 人狼知能プロジェクトと AI

3.1 人狼知能プロジェクトとは

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

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 プレーヤはおしなべて強いプレーヤとなる.しかしなが ら、BBS 人狼で強いプレーヤでも対面でやると、しば し、共同研究プロジェクトである「人狼知能プロジェク しば論理的に正しくても勝つことができない。それを単

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. エンターテイメントとしての魅せる人狼 ★プライン型人狼や、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 ある. 人狼知能プロジェクトでは、人狼知能プロトコル 狼がこれまでに登場している。本章では、エンターテイ 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 広げながら、人狼をプレイしそれをエンターテイメント 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.

人狼知能プロジェクト 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, 各プレーヤの発言内容自体と,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 (ゲーム上の昼の発話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」のプレーヤにとってあな

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

話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」のブレーヤが人狼で

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

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

プレーヤは、客観的視点での情報非決定性により、自

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

かどうかはゲーム終了までわからない (図2).

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.

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 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

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

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

単に役者が人狼を行うのでもない,全く新しい形のエン ンゲームではゲームの勝敗はブレーヤ同士の情報交換に

69 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、○○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

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、ブレーヤは

どのように自らの意思を決定し、選択を行うべきなのだ

発話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 行わなければならず、それはプレーヤ自身にしか決定で

human!

§3 推理:他者の意図のモデル化

なることもある.

ろうか.

きない (図3).

図2 客観的視点での情報非決定性の例. あるプレーヤ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

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 70

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

人 工 知 能 30巻1号(2015年1月)

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

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

ar h

としてきた.

図4 他者から見た自己のモデル化の例.

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

で発言する必要がある

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

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

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, 現実に存在する諸問題

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

これまで、人工知能の標準問題としてさまざまな問題

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、共有するべき大き 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 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 標とした人工知能研究のことである。この汎用人工知能 意見を通すためには、自分の意見が信頼に足るものであ 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 ることを他者に説明する必要がある.この段階では,他 な定義は存在しないが、Adams らにより、汎用人工知 者のモデル化だけではなく、他者から見た自己のモデル 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 化、信頼できる発話の演出が重要となってくる(図4). [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

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

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

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

がある。例としては、将棋など手数を読み切れる終盤に

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

は、人間を交えたゲームとして準備をしている.

ど 15 の能力エリアを設けている.

能力が必要となる [Adams 11].

が提案されてきている。代表的なものとしては、チェス おける完全情報ゲームなどがあげられている [伊藤 14].

人狼知能プロジェクト 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

また, 人狼知能プロジェクトに関連する研究において,

プレーヤの残り数が5人もしくは7人のとき、人

●雑談すると殺される (人狼 BBS) [稲葉 13]

[鳥海 14a]

72

視野に入っている.

狼プレーヤは誰も襲撃しないほうがよい (人狼 AI)

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みてほしい.

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ブ 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]. ブレーヤは 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、 $10 \sim 20$ 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

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ータ 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 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 年はかかるといわれるが、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

3.6 人狼知能プロジェクトにおけるイベントなど

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、人狼知能ブロ

based Werewolf (2013). http://www.aiwolf.org/

[人狼読本 13] 人狼読本, KADOKAWA enterbrain (2013)

http://www.fujitv.co.jp/jinroh/

[人狼 Online 14] 人狼 Online (2014). http://jinrou-online.

[人狼 TLPT] 人狼ザ・ライブプレイングシアター. http://7th-

[人狼~嘘つきは誰だ?~] 人狼~嘘つきは誰だ?~, フジテレビ.

71 実は私は占い師ではないです

話を整理したいのですが

狂人しかいない人狼とか胸が熱いな

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 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,

大学教育において, 例えば, 工学部の授業においては,

学生はま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

人 工 知 能 30巻1号(2015年1月)

自分で判断することが鍛えられる.

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 [稲葉 13] 稲葉通将, 大畠菜央実, 鳥海不二夫, 高橋健一: 雑談ば かりしてると殺される―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―, 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

http://parkl.wakwak.com/~aal/okami/

[CEADEC 14] Computer Entertainment Developers Conference $(CEDEC\ 2014)$ (2014). http://cedec.cesa.or.jp/

[ダンガンロンバ] 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 ダンガンロンバ1・2 超高校級の人狼,

http://www.algernonproduct.co.jp/werewolf/dr1_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

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

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・ 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,

著 者 紹 介・

人狼知能プロジェクト

castle.com/jinrou/

2014年11月14日 受理

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

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 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

73

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,エンターテイメントコン ビューティング, 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、

松原

能研究のすすめ」(2014). http://www.tkomat-lab.com/ [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 エージェントの合同シンポジウム JAWS 2013 [JAWS 開発, Game Programing Workshop, P-9 (2014) 13] および JAWS 2014 [JAWS 14] では, イベントとし [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 て、人狼知能プロジェクトによって、人狼大会の運営と る最適戦略の抽出,情処第 76 回全国大会(2014) 人狼クイズ大会運営を行ってきた.今年度は,人狼知能 [Kobayashi 14] Kobayashi, Y., Osawa, H., Inaba, M., Shinoda, K., Toriumi, F. and Katagami, D.: Development of werewolf match

い師が、実際には占っていない村人Aに対して、「Aは いる.人狼による対話教育も、人狼知能プロジェクトの

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 プレーヤを構築してくれる人が一 情報がプレイヤーに与える影響について, JAWS2013 (2013) 人でも増えたら,嬉しい限りである.人狼ブロジェクト [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 にご興味をもたれた方は、プロジェクト HP(http:// プロジェクト,デジタルゲーム学会 2014(2014) www.aiwolf.org/) をご覧いただければ幸いである. [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) ◇ 参 考 文 献 ◇ [4Gamers] 将棋でプロを倒した AI が、次に狙うのは人狼. 「人狼 [Adams 11] Adams, S. S., et al.: Mapping the Landscape of 知能」プロジェクトに迫る, CEDEC 2014 (2014). http:// Human-Level Artificial General Intelligence, AAAI (2011) www.4gamer.net/games/999/G999905/20140905115/ [aduma 03] 「汝は人狼なりや?」と小一時間問い詰めたい(2003). [人狼 BBS] 人狼 BBS (ブラウザゲーム). http://ninjinix.com/ http://web.archive.org/web/20031123075546/ [人狼知能プロジェクト13] 人狼知能プロジェクトArtificial Intelligence

会員.



報科学研究科助手, 2007年同助教, 2012年東京大 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.

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