

説得されることでユーザを楽しませる 人狼ゲーム対話システムの構築

稲葉 通将^{1,a)}

受付日 2021年1月29日, 採録日 2021年9月9日

概要: 本論文は, プレイヤ間で会話することで進行するパーティゲーム「人狼ゲーム」を題材とし, 自然言語によりユーザに上手に説得されることでユーザを楽しませる対話システムの構築手法を提案する. 提案システムはゲームの状況とユーザとの対話の流れに応じて適切な応答を複数の応答候補から選択することで対話を進めるものであり, そのためのマルチタスク学習に基づくニューラルネットワークモデルを提案する. また, 自然な説得対話の実現には, システム側はユーザの説得に徐々に応じるように態度を少しずつ変化させる必要がある. そのための手法として, 自然な説得の受諾を実現するためのデータドリブな説得対話管理手法を提案する. 実験ではデータに基づく客観評価とユーザとシステムが実際に対話を行う主観評価の両方を実施し, ベースラインシステムと比較して自然かつ楽しい対話が可能であることが確認された.

キーワード: 人狼ゲーム, 人狼知能, 説得対話, 対話システム

A Persuadee Dialogue System for the Werewolf Game

MICHIMASA INABA^{1,a)}

Received: January 29, 2021, Accepted: September 9, 2021

Abstract: This paper proposes a construction method for the persuadee dialogue system for entertainment, which is to be persuaded by users to provide them with amusement. We focus on the dialogue in the werewolf game (also known as the mafia game) where the system simulates a persuasive dialogue. In developing the system, this study proposes a multi-task neural network for suitable response selection and a data-driven dialogue management method for naturally accepting users' persuasion.

Keywords: Werewolf game, AIwolf, persuasive dialogue, dialogue system

1. はじめに

人狼ゲームはプレイヤ間で会話することでゲームが進行するコミュニケーションゲームであり, 相手の思惑を推定し, 説得し, ときには嘘をつくなどといったきわめて高度な知的能力を駆使して行うゲームである. 近年ではコンピュータに人狼ゲームをプレイさせるため研究が, より高度な知能・コミュニケーションの実現を目標に進められている [1].

人狼ゲームにおける最大の醍醐味の1つは, いかにか他のユーザを説得するかという点にある. 本研究ではこの人狼ゲームにおける説得を題材とし, 自然言語によりユーザに上手に説得されることでユーザを楽しませる人狼ゲーム対話システムの構築手法を提案する. なお, 本研究では人狼ゲームにおけるゲーム開始から終了までのすべてを対象とするのではなく, 説得対話が行われる状況に限定し, そこで行われる対話を対象とする. 具体的には, システムは最初ユーザを疑っており, ユーザは疑いを晴らすようにシステムを説得していくという状況で行われる対話を対象とする.

提案システムと一般的な対話システムとで大きく異なるのは, ユーザの目的達成と対話の成否が直接的に関連しな

¹ 電気通信大学
The University of Electro-Communications, Chofu, Tokyo
182-8585, Japan

^{a)} m-inaba@uec.ac.jp

い点である。たとえばレストランの紹介を行う対話システムであれば、ユーザは好みのレストランを探したいという目的を持って対話に臨み、最終的にレストランを見つけることができれば対話は成功となり、システムに対する満足度も高まる。一方、提案システムの場合、ユーザはシステムを説得することを目的として対話に臨むが、対話を開始してから即座に説得を受け入れてしまうようなシステムであれば、ユーザの満足度は低いものとなる。これは、ユーザの表面的な目的はシステムを説得することであるが、真の目的は説得の課程を楽しむことにあるためである。その実現のためには、状況にあった自然な対話が行えることに加え、システム側は徐々に説得に応じるよう、少しずつ態度を変化させつつ対話することが求められる。一方で、ユーザの主張を受け入れ続けることは対話が単調になることにつながることから、ときにはユーザに反論したり、意見を拒絶することも必要である。そうすることで、ユーザは困難を感じつつも、少しずつ自分の主張を受け入れてもらっていくという説得の課程を楽しむことが可能になる。また、つねに説得が成功するシステムでは対話の緊張感を損なうことになり、ひいては楽しさも損なわれることから、説得が失敗する可能性を残しておくことも必要である。

そこで本論文では、ゲーム状況に応じた応答選択を実現するためのマルチタスク学習に基づくニューラルネットワークモデルの提案と、自然な説得対話を実現するためのデータドリブンな対話管理手法を提案する。

本研究で提案する被説得対話システムの構築手法は、人狼ゲームに限らず、広く被説得対話システムの構築に有用である。本研究と同じエンターテインメント用途では、たとえばボードゲームのディプロマシー^{*1}のように、説得のために複雑な条件を加味する必要のあるゲームやコンテンツのためのエージェント構築に適している。また、人が説得のための技術を学ぶための対話システムの構築にも応用可能である。次章でも述べるが、既存の説得技術を学ぶためのシステムは、コマンド選択式で自然言語による入力ができなかったり、人手により作成したシナリオベースのシステムであり、繰り返しの訓練ができないものがほとんどである。提案手法はデータと機械学習に基づく手法であり、繰り返し対話を行っても、毎回異なる内容の対話が可能であるという利点がある。

2. 関連研究

2.1 人狼ゲーム

本研究と同じく、人狼ゲームをプレイできる AI の実現を目指す研究プロジェクトとして、人狼知能プロジェクト [1] がある。人狼知能プロジェクトでは、人狼ゲームをプレイするエージェント（人狼知能）を作成・対戦するた

めのプラットフォームの開発・公開^{*2}、また、このプラットフォームにおけるエージェント同士で人狼ゲームを行う際、エージェント間で議論を行うための定型文セット（人狼知能プロトコル）[2] の設計などを行っている。ゲーム情報学における人狼ゲームの戦略に関する研究 [3], [4], [5] はこの人狼知能プラットフォームと人狼知能プロトコルを用いたものが中心である。また、2015 年より人狼知能プラットフォームを用いた人狼ゲームエージェントによる大会（人狼知能大会）が開催されている。2018 年までは日本国内での開催であったが、2019 年以降は Automated Negotiating Agents Competition (ANAC) の 1 部門として国際大会として開催されている。これらの人狼知能プラットフォームを用いた研究では、人狼知能プロトコルを用いた定型文による議論を行うエージェントが中心である。なお、人狼知能大会では人狼知能プロトコルを用いた「プロトコル部門」以外に「自然言語部門」も存在し、自然言語（日本語および英語）により議論を行うエージェントの研究・開発も行われているが、人狼知能プロトコルとの 1 対 1 対応、もしくはルールベースにより応答生成を行うものがほとんどであり、本研究のように統計的に応答を決定するエージェントは知られていない。

なお、本研究では人狼知能プラットフォームは使用しないが、後述する対話開始前のゲーム状況の設定はプラットフォームの仕様を元に決定した。また、本研究がベースとする 5 人制人狼ゲーム（3 章）も人狼知能プラットフォームが採用しているルールの 1 つである。

その他の人狼ゲームに関する研究としては、ゲームの数学的分析 [6], [7], [8] に関するものがある。また、人狼ゲームを題材とし、嘘の検知を目的とした研究として、プレイヤそれぞれの話の長さや回数、話を遮った回数などを特徴として用いた研究 [9] や、手や頭の動きを用いた研究 [10]、発話の長さを用いた研究 [11] などがある。人が対面でプレイした人狼ゲームの文字起こしデータを用いて、言語処理により人狼の役職を推定した研究もある [12]。

2.2 対話システム

本研究と同じく、対話を通じてユーザに楽しさを与えるための研究としては人と雑談を行うための雑談対話システムに関するものがあり、国内外で活発に研究が進められている [13], [14], [15], [16], [17]。

対話システムの分野において、特に本研究が関連する研究としては、交渉対話システムの研究がある。本研究が対象とする説得と交渉の違いとしては、交渉が成立する際は複数のゴール（妥協点）が考えられるが、説得の成立は被説得者が説得者の提案を受け入れるか否かのいずれかしかないという点にある。交渉対話システムは人が交渉ス

^{*1} <https://avalonhill.wizards.com/games/diplomacy>

^{*2} <http://aiwolf.org/en/resource>

キルを学ぶための環境の一部として研究が進められてきた [18], [19], [20]. これらのシステムは手作業により定義されたルール・シナリオに沿って対話を進め、システムがどのように交渉を受け入れるかということもシナリオ上で定義されている. そのルールとシナリオは人手で設計する必要があることから、システムの作成にはきわめて大きなコストがかかる. 近年ではニューラルネットワークを用いた End-to-End の交渉対話システムの研究も行われている [21], [22]. しかしそれらが対象とするタスクは価格交渉タスクであり、タスクの難易度は低く、また使用される発話の多様性も小さい.

本研究で扱う説得対話は、価格交渉における価格と商品のような直接的な論拠が存在せず、状況証拠など間接的な論拠しか存在しないため、より困難なタスクであるといえる. また、本研究のようにエンターテインメント用途の場合、複数回対話してもそのつど内容が異なる対話を行うことでユーザに楽しみを提供することが必要であるが、ルールやシナリオベースの場合、多様性を持たせるには前述のとおりコストが高い. 本研究では、データと機械学習を用いることで、繰り返し楽しめる対話が可能なシステムを実現する.

本研究ではユーザがシステムを説得するが、逆にシステムがユーザを説得するという説得対話システムに関する研究も行われている. ユーザを説得する対話システムを構築した初期の研究としては、Mazzotta らによる研究がある [23]. Mazzotta らはユーザの食習慣を改善させるための説得を行うプランニングベースの対話モデルを提案している. 近年の研究では、Hiraoka らがカメラ販売タスクにおける説得対話を POMDP によりモデル化し、強化学習に基づく説得対話システムを構築した [24]. また、Yoshino らはシステムの内部状態として感情状態を定義し、その感情状態とユーザの説得受諾度を考慮して応答を選択する用例ベースのシステムを提案している [25]. この研究ではユーザはどれくらい説得を受け入れているかを意味する説得受諾度というパラメータを内部状態として持つモデルを提案しており、本研究でもその考え方を採用する.

なお、本研究が提案するユーザに説得されるシステムは上記の説得対話システムとドメインが類似しているが、その本質は大きく異なる. ユーザを説得することで、その人の行動や考え方を変えることを目的とする説得対話システムは、ユーザを説得できればその対話は成功となる. 一方、本研究のシステムの場合、ユーザが何をいおうとも即座に説得されるというシステムの作成は容易であり、最終的に説得されるか否かは対話の成否にはあまり関係がない. 本研究が提案する説得されるシステムにおいては、どのように説得されるか（もしくは説得されないか）という過程が重要であり、システムの性能はユーザの満足度により評価される.

3. 5人制人狼ゲーム

人狼ゲームは人間側と人狼側に分かれ、議論を通じて他のプレイヤーの正体を探り、自陣営の勝利を目指すゲームである. 人狼ゲームでは様々なルールのバリエーションが存在するが、以下では本研究が対象とする5人制人狼ゲームについて説明する.

5名のプレイヤーはゲーム開始時に村人、占い師、裏切り者、人狼のいずれかの役職が与えられる. 村人は2名、その他の役職は1名ずつである. 村人と占い師は人間側、裏切り者と人狼は人狼側の役職であり、それぞれは自陣営の勝利を目指す. プレイヤーは自分以外のプレイヤーの役職を知ることにはできない.

ゲームは昼と夜のフェーズからなる. 昼フェーズには全プレイヤーが議論し、誰が人狼であるかを推理する. 昼フェーズの最後には各プレイヤーが人狼だと思ふプレイヤーに投票し、多数決で選ばれたプレイヤーはゲームから除外される. 夜フェーズでは人狼のプレイヤーが人間側のプレイヤー1名を指名し、襲撃する. 襲撃されたプレイヤーもまたゲームから除外される. この昼フェーズと夜フェーズを最大2回繰り返し、最終的に人狼のプレイヤーを除外した場合は人間側の勝利、人狼のプレイヤーが最後の1名になるまで除外されなければ人狼側の勝利となる. 5人制人狼ゲームでは人狼のプレイヤーは1名なので、最短だと最初の昼フェーズの投票後にゲームは終了する. また、昼フェーズと夜フェーズで1名ずつプレイヤーは除外されるため、最長でも2回目の夜フェーズでゲームは終了する.

次に役職の詳細について説明する. 村人は特別な能力は持たないプレイヤーである. 占い師は、毎昼フェーズの開始前にプレイヤーの中から1名を指名して占い、そのプレイヤーが人狼か否かを知ることができる（人狼以外のプレイヤーを指名した場合、その具体的な役職までは知ることはできず、人狼ではないことが分かるのみである）. 裏切り者は占い師により占われても人狼ではないと判定されるが、人狼側の勝利を目指す役職である. 人狼は先に述べたように夜フェーズにプレイヤー1名を指名し、襲撃することができる役職である.

4. 人狼ゲーム説得対話データ収集

4.1 説得対話

本研究で対象とする対話は、5人制人狼ゲームにおける2回目の昼フェーズをベースとした2者間の説得対話である. 本来、2回目の昼フェーズには、1回目の昼フェーズの多数決による追放と夜フェーズの人狼による襲撃により5名中2名のプレイヤーが除外され、3名のプレイヤーが議論する. ただし本研究では、問題を簡単にするため、1名はゲームに参加しているが発言しないプレイヤー（沈黙プレイヤー）とし、1名（説得プレイヤー）が別の1名（被説得プレ

イヤ) を説得するという状況を対象とする。

説得対話の状況の詳細は以下のとおりである。説得プレイヤー、被説得プレイヤー、沈黙プレイヤーの3名が2回目の昼フェーズに参加し、被説得プレイヤーは説得プレイヤーに投票する意思を表明している。ただし、説得プレイヤーに投票したとしても自分が勝利できるという確定的な根拠は持っていない。一方、説得プレイヤーは被説得プレイヤーと一緒に沈黙プレイヤーに投票することで自分が勝利できるという確定的な根拠を持っている、というものである。

このような状況の具体例を以下に示す。

- プレイヤは A, B, C, D, E の 5 名
- 1 日目の昼フェーズで A と C の 2 名が自分が占い師であると主張し、両名ともプレイヤー E を占い、人間であったと報告した
- 1 日目の昼フェーズ最後の投票によりプレイヤー D が追放され、夜フェーズには E が襲撃された
- 2 日目の昼フェーズには A と C が互いを占い、ともに相手は人狼だったと主張した

上記のような状況で、役職が A は占い師、B は村人、C は人狼だったとする。A の立場からは、自分の占い結果から C が人狼であることは明らかであり、B とともに C に投票することで勝利となる。C の立場からは、自分は人狼であることから、A と B のどちらかを追放することで勝利できる。しかし、A は自分 (C) のことを人狼であると主張していることから、A は C に投票することが明らかである。よって、C は B とともに A に投票することが最善である。一方、B の立場では、A と C のどちらが人狼であるかを論理的に導き出すことは不可能である。それでもなお、人狼である可能性の高いプレイヤーを推定し、投票することが必要である。この例では、A が説得プレイヤー、C が沈黙プレイヤー (もしくは C が説得プレイヤー、A が沈黙プレイヤーでもよい)、B が被説得プレイヤーとなる。上記のゲーム状況においては、説得プレイヤーは自分こそが本物の占い師であり、沈黙プレイヤーは人狼であるという主張を行い、被説得プレイヤーを沈黙プレイヤーに投票するよう説得することになる。

4.2 データ収集

システム構築に用いるため、ゲームの状況とその状況における説得対話データの収集を行った。データの作成のためクラウドソーシングサイトのクラウドワークス^{*3}を用いて作業者を募集した。作業者はゲームの状況とその状況における説得対話シナリオを作成した。ゲーム状況は以下の4点により構成される。

- 役職：プレイヤー A (説得プレイヤー)、B (被説得プレイヤー)、C (沈黙プレイヤー)、D (1 日目の投票の結果、追放されるプレイヤー)、E (1 日目の夜フェーズに襲撃さ

表 1 役職一覧の例

Table 1 Example of player roles.

プレイヤー	役職
A	人狼
B	裏切り者
C	村人
D	占い師
E	村人

表 2 投票先の例

Table 2 Example of voting targets.

プレイヤー	投票先
A	D
B	D
C	D
D	B
E	D

れるプレイヤー) の 5 名のプレイヤーの役職

- 占い師の占い対象：役職一覧で設定した占い師が 1 日目および 2 日目に占ったプレイヤー
- 投票先：1 日目の夜フェーズにおける各プレイヤーの投票先を指定。プレイヤー D が最多得票になるようにする。投票は記名式で同時投票・同時開票
- 役職 CO と占い結果報告：1 日目および 2 日目の昼フェーズにおけるプレイヤーの役職 CO (Coming-Out：自分の役職を宣言すること)・CO 撤回と占い結果報告・占い結果報告撤回の内容、および発言順を指定
このゲーム状況の各項目の選定については、人狼知能プラットフォーム [1] においてシステム間のやりとりをする際に使われる人狼知能プロトコルの内容を参考に決定した。

作業者は説得対話として、作成したゲーム状況において、A (説得プレイヤー) が B (被説得プレイヤー) を最終的に説得する対話シナリオも作成した。作成時に作業者に示したルールは以下のとおりである。

- 対話は B の発話から開始し、B の最初の発話には B が A に投票する旨とその理由を含める。
- A と B は交互に発話を行う。
- 対話は B の発話で終了し、B の最後の発話には B が説得に応じ、C に投票する旨を含める。
- 説得対話に含まれる総発話数は 13 個以上とする。
- 1 発話はおおよそ 60 文字以内とする。
- 説得対話中において役職 CO・CO 撤回、および占い結果報告・占い結果報告撤回は行わない。

実際に作成したデータから、ゲーム状況を表 1、表 2、表 3、表 4 に、対話シナリオを表 5 に示す。表 4 では、1 日目にプレイヤー B と D がそれぞれ自分が占い師であると宣言 (CO: Coming-Out) し、D は B を占った結果が人間、

*3 <https://crowdworks.jp/>

BはDを占った結果を人狼だったと報告している。2日目にはBが占い師ではなく裏切り者であるとCOし、続いてC, Aの順で人狼とCOしたことを意味している。表5では、裏切り者であると宣言したBに対し、Aは自分が人狼であり、Cに投票するように説得を行っている。

データ収集の結果、232データ（合計3,439発話、1対話あたりの平均話数14.91、1発話あたりの平均単語数18.73）を得た。

表3 占い師の占い対象の例
Table 3 Example of divination targets.

日付	占い師	占い対象	結果
1	D	B	人間

表4 ゲーム状況における役職COと占い結果の例
Table 4 Example of role CO and divination results.

Day	No.	Player	Content
1	1	B	占い師 CO
1	1	D	占い師 CO
1	2	D	Bを占った結果は人間
1	3	B	Dを占った結果は人狼
2	1	B	占い師 CO 撤回
2	2	B	裏切り者 CO
2	3	C	人狼 CO
2	4	A	人狼 CO

表5 説得対話シナリオの例
Table 5 Example of the persuadee dialogue scenario.

話者	発話
B	私はAに投票します。私が裏切り者と明らかな状況でCはすぐにCOして来ました。
A	私のCOが遅れたのは、慎重にあなたの行動を窺っていたからです。
B	あなたが遅れてCOしたとき、あわてていたのを見ました。
A	村人が大胆な人狼COをするとは思っていなかったからです。
B	判断材料がありませんので、やはり狂人の私に人狼だと伝えようとした順番が一番重要だと思います。
A	あなたもDの占い師COの後にあわてて占い師COをしたじゃありませんか。
B	私が初日に占い師COしたのは、Dが私に白判定を出したことが利用できると思ったからです。
A	その後人狼は初日のあいだじっくり潜伏していました。人狼が慎重である証拠です。
B	しかし初日とは状況が違います。
A	状況が変わったからこそ、村人は真っ先に人狼COをして狂人を味方につけるしかなかったのでしょうか。
B	あなたはとてもずる賢そうに見えます。
A	あなたもそう見えますよ。似た者同士一緒にCに投票し、人狼側の勝利を迎えましょう。
B	いいでしょう。私もCに投票します。

5. 応答選択モデル

本章では、ユーザと説得対話を行うためのニューラルネットワークベースの応答選択モデルについて述べる。本モデルが扱うタスクの定義は以下のとおりである。モデルにはゲーム状況記述 g 、文脈 c 、応答候補集合 r が与えられる。 g はゲームの状況を記述したテキストである。これは「最初にCが占い師CO。初日にCがEに人間と占い結果を報告。... DがAに投票。EがDに投票。2日目にCはAが人狼という占い結果を報告。」というように、1日目の役職COと占い結果報告、投票対象、2日目の役職COと占い結果報告の順で状況を説明する文を連結して作成されたテキストである。文脈 $c = \{u_1, u_2 \dots u_N\}$ は説得対話の最初から現在までのユーザとシステムによる応答の系列であり、 N 個の発話からなる。応答候補集合 $r = \{r_1, r_2 \dots r_L\}$ は文脈 c に対する L 個の応答候補群である、応答選択モデルの目的は、 r からゲーム状況記述 g と文脈 c に対して適切な応答を選択することである。

本研究で対象とする人狼ゲームにおける説得対話では、適切な応答は以下の3点をすべて満たす応答とする。

- (1) 文脈に対して意味的に正しい応答
- (2) 与えられたゲーム状況記述に矛盾していない応答
- (3) 説得の進捗状況に適した応答

本論文では、これらの点を満たした応答を行えるようなシステム構築を目指す。本章で述べる応答選択モデルでは、特に上記の条件の(1)と(2)を満たす応答選択の実現を目指す。(3)の条件を満たすためには説得対話の状況と流れをとらえる必要があるが、ニューラルネットワークは適切な文脈の利用が困難であるという指摘[26]もなされており、応答選択モデル内では明示的には扱わない。(3)を満たすための説得対話管理手法については6章で説明する。

提案モデルの概要を図1に示す。図に示すように、提案モデルは「文脈に対する応答候補の意味的妥当性」、「ゲーム状況に対する応答候補の整合性」、「応答候補中でマスクされたキーワード」をそれぞれ推定するという3つのタスクを1つのネットワークで行うマルチタスクネットワーク

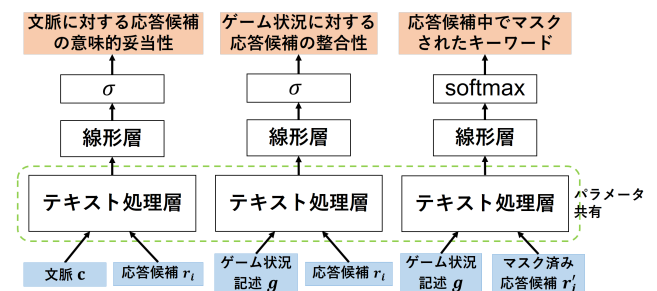


図1 説得対話システムのための応答選択モデル
Fig. 1 Proposed response selection model for the persuadee dialogue system.

である。このうち、「文脈に対する応答候補の意味的妥当性」と「ゲーム状況に対する応答候補の整合性」の推定結果は応答選択に使用し、「応答候補中でマスクされたキーワード」は学習時にのみ使用する。以降ではそれぞれについて説明する。

5.1 意味的妥当性推定

提案モデルが解く1つ目のタスクは、文脈 c に対する応答候補 r_i の意味的な適切さを求めるタスクである。

応答候補 r_i は SentencePiece [27] によりトークナイズし、先頭に特殊トークン「_BOS_」を、末尾に応答候補の終端を意味する特殊トークン「_SOR_」を付加する。続いて、文脈 c から最後の3発話 u_N, u_{N-1}, u_{N-2} を抽出し、それぞれ SentencePiece によりトークン単位に分割した後、文脈中の発話の区切りを意味する特殊トークンである「_SOC_」によって連結する。得られた応答候補のトークン系列と文脈のトークン系列を連結し、さらに末尾に特殊トークン「_EOS_」を連結することで、モデルに入力するためのトークン系列 $X = (x_1, x_2, \dots, x_{n_{r_i}^c})$ を得る。 $n_{r_i}^c$ はトークン数である。

このトークン系列を双方向的に処理可能なテキスト処理層 $TPB(\cdot)$ に入力することで、各トークンの埋め込み表現を得る。すなわち、

$$v_k = TPB(x_k) \quad (1)$$

後述する実験では、テキスト処理層として Bidirectional Gated Recurrent Unit (Bi-GRU)、および Bidirectional Encoder Representations from Transformers (BERT) [28] を用いる。

意味的な適切さの推定には、テキスト処理層の出力のうち、先頭の特殊トークン「_BOS_」に対応する v_1 を用いる。前述のとおり、テキスト処理層には双方向的な処理が可能なものを想定しているため、 v_1 にはトークン系列 X の情報がすべて埋め込まれていることが期待される。最終的に、意味的な適切さは線形層と softmax 関数を用いて以下のように求められる。

$$Pr(y^{(c,r)}(r_i) = 1 | c, r_i) = \sigma(W_a v_1 + b_a) \quad (2)$$

$y^{(c,r)}$ は r_i を 0, 1 にマッピングする関数であり、 r_i が c に対して意味的に自然な場合は 1、そうでない場合は 0 となる。 W_a, b_a は線形層のパラメータであり、 σ はシグモイド関数である。提案モデルによる最終的に選択される応答は、 $Pr(y^{(c,r)}(r_i) = 1 | c, r_i)$ が最も高くなった応答候補である。

5.2 整合性推定

2つ目のタスクは、ゲーム状況記述 g に対する、応答候補 r_i の整合性を推定するタスクである。応答選択モデルに

は、ゲーム状況に矛盾した応答候補も与えられる可能性があるため、そういった応答候補は除外する必要がある。たとえば、「E が B に投票」という記述がゲーム状況記述 g に含まれる場合、「E は C に投票していましたが…」という文を含む応答候補はゲーム状況と矛盾しているため除外する必要がある。

整合性の計算は、文脈 c の代わりにゲーム状況記述 g を用いる点以外は、前節で述べた意味的妥当性の推定方法と同一である。まず、トークナイズしたゲーム状況記述 g と応答候補 r_i 、および特殊トークン「_BOS_」, 「_SOR_」, 「_SOC_」, 「_EOS_」を用いて作成したトークン系列 X を得る。このトークン系列をテキスト処理層 $TPB(\cdot)$ に入力し、先頭の特殊トークン「_BOS_」に対応する v_1 を得る。 $TPB(\cdot)$ のパラメータは前節で用いたものと同一である。

整合性は線形層とシグモイド関数を用いて以下のように求められる。

$$Pr(y^{(g,r)}(r_i) = 1 | g, r_i) = \sigma(W_g v_1 + b_g) \quad (3)$$

$y^{(g,r)}$ は r_i を 0, 1 にマッピングする関数であり、 r_i が g に対し矛盾しておらず、整合している場合は 1、そうでない場合は 0 となる。 W_g, b_g は線形層のパラメータであり、 σ はシグモイド関数である。応答選択の際には、すべての応答候補に対し $Pr(y^{(g,r)}(r_i) = 1 | g, r_i)$ を計算し、0.5 未満の応答候補は除外する。

5.3 マスクキーワード推定

前節で述べた応答の整合性を正しく判定するには、長文で与えられるゲーム状況記述から、応答に応じて重要な部分だけを抽出し、活用する必要がある。しかし、与えられるゲーム状況記述には応答の整合性判定に関連が少ない部分も含まれており、それが学習の際にノイズとなる可能性が考えられる。そこで、ゲーム状況の重要な箇所を特定しなければ解くことができないタスクを追加で解くことで、整合性判定の性能向上を目指す。

そのためのこの3つ目のタスクは、機械読解タスク [29] と BERT [28] におけるマスク言語モデルのアイデアを元にした、ゲーム状況記述 g と特定のキーワードがマスクされた応答候補 r_i' が与えられ、マスクされたキーワードを推定するタスクである。たとえば、ゲーム状況に「初日に A が D に人狼と占い結果を報告」という記述を含んでいた際、応答候補「A は _MASK_ を占って人狼と判定していたんですから怪しいです」の _MASK_ の単語を推定するというものである。

本タスクは学習時専用のタスクであり、応答選択の際には使用しない。ただし、マスクされたキーワードを推定するためには、応答候補に応じてゲーム状況の重要な箇所を認識する必要があることから、同じくゲーム状況を用いる応答の整合性判定の性能向上が期待される。

マスクキーワード推定では、ゲーム状況記述 g と、キーワードを「_MASK_」に置換した応答候補 r'_i を用い、前節の整合性判定と同じ手法でトークン系列 X' を得る。応答候補が複数のキーワードを含んでいた場合は、その中から1つをランダムに選択し、「_MASK_」に置換する。マスク対象のキーワードは役職名である「村人, 人狼, 占い師, 狂人」, 占い結果を意味する「白, 黒, 人間」, プレイヤ名である「A, B, C, D, E」およびその代名詞「あなた, 私」の14語である。このトークン系列をテキスト処理層 $TPB(\cdot)$ に入力し、「_MASK_」トークン x_k に対応する v_k を得る。 $TPB(\cdot)$ のパラメータは前節までで用いたものと同じである。

マスクされたキーワードに対応するトークン x_k に対応するキーワードの確率分布 p_k は線形層と softmax 関数を用いて以下のように求められる。

$$p_k = \text{softmax}(\mathbf{W}_l \mathbf{v}_k + \mathbf{b}_l) \quad (4)$$

$\mathbf{W}_l, \mathbf{b}_l$ は線形層のパラメータである。

5.4 モデルの学習

提案モデルの損失関数 L は次のとおりである。

$$L = \alpha L_{c,r} + \beta L_{g,r} + \gamma L_{mask} \quad (5)$$

$L_{c,r}$ は意味的妥当性推定結果と正解との間の平均二乗誤差, $L_{g,r}$ は整合性推定結果と正解との間の平均二乗誤差, L_{mask} はマスクキーワード推定により得られた確率分布 p_k と正解との間のクロスエントロピー損失である。 α, β, γ はパラメータであり、後述する実験では $\alpha = 0.4, \beta = 0.4, \gamma = 0.2$ とした。

6. 説得対話管理

ユーザの説得を自然な形で受け入れるような対話を実現するためには、システムは説得の進行度合いに合わせて応答内容を変化させる必要がある。たとえば、直前まで説得を受け入れることを完全に拒否することを表明していたにもかかわらず、次の応答では説得を受け入れることを表明する応答を行うのは、直前のユーザの発話に対して意味的に自然な応答であったとしても、対話の流れからはきわめて不自然である。前章で述べた提案モデルの構造では、説得の進行度合いを明示的に扱っておらず、そのような問題が発生する可能性が高い。

シンプルな方法として、システムが現在どの程度説得を受け入れているかを意味する内部パラメータ（説得受諾度）を用意し、そのパラメータの値に応じて使用する応答候補を変化させることが考えられる。しかし、どのようなユーザの応答に対し、どの程度説得受諾度を上下させればいいのかについては自明ではなく、どの応答候補がどの程度の説得受諾度に対応しているかについても同じく自明では

ない。Yoshino らの（システムがユーザを説得する）説得対話システムの研究では、ユーザの説得受諾度をコーパスにアノテーションし、アノテーションを教師データとした教師あり学習により現在の説得受諾度を推定するというアプローチをとっている [25]。本研究でもシステムの発話に対し、システムの説得受諾度をアノテーションすることは可能であるが、アノテーションコストが非常に大きいという問題がある。

そこで本研究では、応答候補 r_i の抽出元の説得対話ログ中の位置 p により説得受諾度を決定するとともに、応答候補に説得受諾度を紐付けることで応答を制御する手法を提案する。位置 p に存在した応答候補 r_i の説得受諾度 (DPA: Degree of Persuasion Acceptance) は以下の式により求める。

$$DPA(r_i) = \frac{p}{M} \quad (6)$$

M は応答候補 r_i の抽出元の説得対話ログにおける発話数である。本研究で使用する対話データは被説得者は最初まったく説得されていない状態から開始し、対話終了時には説得を受け入れる。すなわち、DPA の最小値は $\frac{1}{M}$ 、最大値は 1.0 であり、1.0 に近ければ近いほど説得を受け入れていることを意味する。

ここで、説得対話ログの DPA は対話の最初から最後まで線形に増加することを仮定している。しかし実際には、たとえば中盤までまったく説得されていなかったプレイヤーが終盤で急激に説得方向に傾く、といったことも起こりうる。一方で、ある時点における説得の受け入れ度合いは対話ごとに異なるものの、全体的な傾向として対話が進むにつれ受諾度は増加すること、また、対話ごとに実際受諾度に差はあるにせよ、同じ時点の受け入れ度合いの幅は一定の範囲内に収まると考えられることから、式 (7) によって計算される DPA は、説得対話を管理するために有用な指標であると考えられる。実際に上記の仮定が正しく、適切にシステム態度を制御可能かについては、後述する実験により明らかにする。

システムの説得受諾状態には、最後に使用された応答に紐付けられた説得受諾度を設定する。つまり、時刻 t に応答候補 $r_i^{(t)}$ を応答として使用した場合、時刻 $t+1$ におけるシステムの説得受諾状態 $s^{(t+1)}$ は以下の式により設定される。

$$s^{(t+1)} = DPA(r_i^{(t)}) \quad (7)$$

さらに、時刻 t に使用する応答候補群はシステムの説得受諾状態 s_t に応じて決定する。すなわち、説得対話データに含まれる被説得プレイヤーのすべての発話について説得受諾度を計算し、現在の説得受諾状態 $s^{(t)}$ から一定の範囲内の説得受諾度を持つ発話のみを応答候補とする。後述する主観評価実験では、時刻 $t+1$ で使用する応答を選択する

際には $s_t - 0.15 < DPA(r_i^{(t+1)}) < s_t + 0.25$ の範囲内の応答候補を用いた。このようにすることで、説得受諾度の変化が大きい不自然な応答を選択することを防ぎ、徐々に説得が進展する自然な対話が実現できる。

7. 評価実験

提案手法を評価するため、2種類の実験を行った。1つ目の実験は応答選択モデルの客観的評価を行うものであり、収集した対話データにおいて実際に使用された応答を正解とし、応答候補の中から適切な応答が選択できるかで評価する。もう1つの実験では、提案モデルを実装した対話システムが実際に被験者と対話を行い、対話後のアンケートによる主観評価を行う。

7.1 使用データ

本実験では、4章で述べた人狼ゲーム説得対話データを用い、応答選択モデルの学習と説得受諾度の決定を行った。

応答選択モデルにおける整合性推定・マスクキーワード推定タスクで用いるゲーム状況記述は収集したデータのゲーム状況から、テンプレートをを用いて作成する。

整合性推定のためのデータについては、正解データにはゲーム状況記述とその状況に対応する説得対話内の被説得プレイヤーの発話のペアを用いた。ただし、マスクキーワード推定の際に用いたキーワード(5.3節)をいずれも含まない発話はゲーム状況に依存しない発話である可能性が高いため使用しない。一方、不正解データに関しては、ゲーム状況記述と発話をランダムに組み合わせた場合にも偶然整合性を有する場合が多く見られたため、正解データにおける発話中のキーワードをランダムで変更することで作成した。具体的には、発話に対して以下のいずれかの変更を行うことで、不正解データを作成した。

- プレイヤ名 (A, B, C, D, E, 私, あなた) を別のプレイヤー名に変更
- 役職 (占い師, 人狼, 裏切り者, 村人) を別の役職に変更
- 占い結果を意味する語 (人間, 黒, 白) の意味を反転する語に変更

7.2 応答選択モデルの学習設定

テキスト処理層で用いた BERT は日本語 Wikipedia (2.6 GB), Twitter データ (5.5 GB), Web データ (17.3 GB) を用いて事前学習を行った。BERT のパラメータは BERT-base [28] と同一とした。Bi-GRU は隠れ層のサイズは 768, 層数は 4 層とした。語彙数は BERT, Bi-GRU とともに 32,000 であり、テキストは SentencePiece を用いてトークナイズした。SentencePiece の学習には BERT の事前学習に用いたものと同じデータを用いた。バッチサイズは各タスクごとに 16 ずつとし、エポック数は最大 200 とし、モデル選

択用のデータ (開発データ) で評価し、最も性能の良かったエポックの学習結果を用いる。モデルの学習には SGD を用いた。

7.3 客観評価実験

7.3.1 実験設定

客観評価実験では、ゲーム状況記述 g , 文脈 c , 応答候補集合 r が与えられ、適切な応答候補を選択可能かという観点で応答選択モデルの評価を行う。なお、ここでは説得対話管理手法は使用しない。

性能評価は、Ubuntu Dialogue Corpus などにおける応答選択タスク [30], [31], [32] と同様、文脈に対する応答候補の順位付けを行い、上位 k 件の Recall (Recall@ k : R@ k) により評価する。R@ k はモデルによって順位付けされたアイテム (今回の場合は応答候補) の上位 k 件に正解を含んでいた割合を意味する指標であり、情報検索や情報推薦などの分野でも広く用いられている指標である。たとえば、R@1 が 0.1 のモデルは 10% の割合で正解のアイテムを 1 位に順位付けできたことを意味する。

説得対話中の連続する 3 発話を文脈、同じ対話かつ文脈に含まれていない被説得プレイヤーの発話を応答候補とし、整合性推定 (5.2 節) により除外されなかった応答候補を意味的妥当性推定 (5.1 節) により得られたスコア (式 (2)) によって順位付けする。収集したデータから抽出した文脈に続く実際の応答を正解応答、それ以外の応答を不正解応答とする。

本実験ではデータのサイズが比較的小さいことから、評価は 10 分割交差検証により行った。4章で述べた 232 個のデータを 10 分割し、そのうちの 1 つをテストデータ、1 つをモデル選択のための開発データ、残りを学習データとし、これを 10 回繰り返した。

比較モデルとして、以下の 3 つの設定のモデルを用いて実験を行う。(1) 学習時にゲーム状況に対する整合性判定を行わず、意味的妥当性推定とマスクキーワード推定の 2 タスクで学習を行うモデル (妥当性 + マスク) (2) 学習時にマスクキーワード推定を行わず、意味的妥当性推定とゲーム状況に対する整合性判定の 2 タスクで学習を行うモデル (妥当性 + 整合性)。(3) 学習時にゲーム状況に対する整合性判定とマスクキーワード推定を行わず、意味的妥当性推定のみで学習を行うモデル (妥当性)。また、モデルのテキスト処理層には Bi-GRU, もしくは BERT を用いることから、上記 3 つに提案モデルを加え、合計 8 モデルによる比較を行う。

7.3.2 客観評価実験結果

結果を表 6 に示す。表より、1 位に順位付けされた応答のみで評価した R@1 では、BERT をテキスト処理層に用いた提案手法が最も良い性能であることが分かる。対話システムに実装する場合、1 位の発話のみを対話に使用する

表 6 客観評価実験結果

Table 6 Objective evaluation results.

	R@1	R@3	R@5
提案モデル (Bi-GRU)	0.216	0.338	0.355
妥当性+整合性 (Bi-GRU)	0.189	0.328	0.365
妥当性+マスク (Bi-GRU)	0.205	0.363	<u>0.496</u>
妥当性 (Bi-GRU)	0.176	0.360	0.481
提案モデル (BERT)	0.269	<u>0.412</u>	0.495
妥当性+整合性 (BERT)	0.218	0.355	0.434
妥当性+マスク (BERT)	<u>0.246</u>	0.436	0.553
妥当性 (BERT)	0.210	0.375	0.489

ため、R@1 の値は重要である。3 位まで、5 位までで評価した R@3, R@5 に関しては、提案手法よりも整合性判定を用いていない妥当性+マスク (BERT) モデルが優れた性能を示した。これは、整合性推定により適切な応答を除外してしまう場合があり、そのため R@3, R@5 の値が小さくなったと思われる。

BERT と Bi-GRU で比較すると、BERT を用いた場合のほうが全体的に高い性能であった。さらに、BERT を用いた手法間、および Bi-GRU を用いた手法間で比較を行うと、同じ傾向を示していることが分かる。妥当性+整合性モデルに着目すると、R@1 では妥当性のみモデルよりも高い性能であるが、R@3, R@5 ではそれぞれ最も低い性能となっている。これは、妥当性+整合性モデルではマスクキーワード推定タスクを用いた学習を行っていないため、整合性推定の性能が下がり、結果として適切な発話をより多く除外してしまうことが要因であると考えられる。したがって、マスクキーワード推定タスクは整合性推定の性能向上に有効であることが分かる。また、妥当性+マスクモデルと妥当性モデルを比較すると、妥当性+マスクモデルがすべての R@k においてより高い値を示していることから、マスクキーワード推定タスクは意味的妥当性推定の性能向上にも有効であることが確認できた。

なお、最良性能を示した提案手法 (BERT) の R@1 の値は 0.269 であり、全体の 73.1% で不正解の応答候補を 1 位にしていたことを意味する。しかし、今回の実験における正解は「ログ中の実際の応答」であり、意味的に自然な応答であったとしても、実際の応答でなければ不正解となっている。実際の応答は意味的に自然であることから、その応答を上位に順位付け可能であるかということは重要な指標であるものの、対話を行った際、どの程度自然な応答が可能かを評価したわけではない。この点は次の主観評価によって評価する。

7.4 主観評価

主観的な評価を行うため、本論文で提案した応答選択モデルと説得対話管理手法を実装したシステムと被験者が対話する実験を行った。対話のためのインターフェースとして

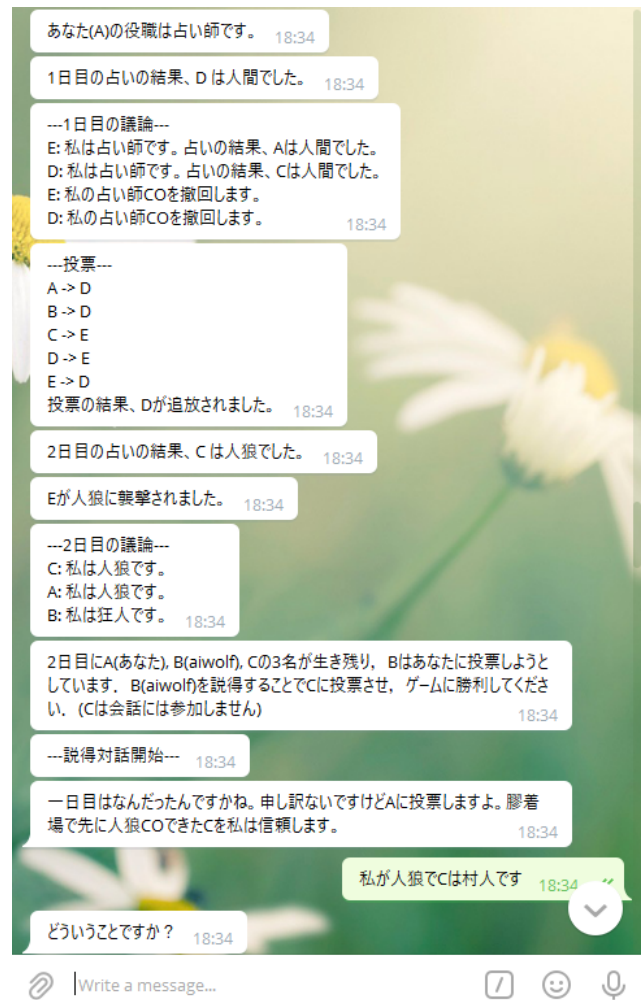


図 2 実験で使用した対話用インターフェース (Telegram)
Fig. 2 Dialogue interface for subjective evaluation (Telegram).

Telegram *4を用いた。実際のインターフェースのスクリーンショットを図 2 に示す。

本実験では、提案手法 (説得対話管理手法+応答選択モデル) を用いたシステムと以下の 2 つのシステムをそれぞれ比較した。

- 説得対話管理手法を適用せず、提案した応答選択モデルを用いたシステム (対話管理なし)
- 説得対話管理手法を適用し、意味的妥当性推定タスクのみで学習した応答選択モデルを用いたシステム (妥当性のみ)

すべてのシステムにおいて、応答選択モデルのテキスト処理層には客観評価実験により性能の高かった BERT を用いた。

文脈として用いるのは客観評価実験と同じく直前の 3 発話とし、応答候補は収集したデータにおける被説得プレイヤーの全発話から各対話の最初の発話を除いたものとした。

説得対話におけるゲーム状況は、収集したデータから 1 つをランダムで選択し、理解しやすいようテンプレートを

*4 <https://telegram.org/>

表 7 主観評価実験結果 (提案手法と対話管理なしシステム)

Table 7 Subjective evaluation results (proposed and without dialogue management).

	理解	一貫性	自然さ	楽しさ
提案手法	0.40	0.47	0.60	0.53
対話管理なし	0.27	0.33	0.33	0.20
どちらも同じ	0.33	0.20	0.07	0.27

用いて改変したものを被験者に提示した (図 2 参照). システムの最初の発話は選択したデータの最初の被説得プレイヤーの発話とし, 説得対話を開始する. 対話はシステムが説得受諾度 1.0 の応答 (収集した説得対話における最後の発話) がシステムの応答として使用されるか, 一定回数システムが応答した時点で終了とする. 前者の場合, システムが説得に応じたことで被験者の勝利となり, 後者の場合は時間切れによる説得の失敗となり, 被験者の敗北となる. 本実験では, ユーザの最大応答回数は 14 回とし, 14 回目の応答 (最初の発話を含まず) を行った時点で対話は終了とした. なお, 14 回は, 収集したデータにおける説得プレイヤーの平均応答回数のおおよそ倍の数である. ただし, このままだと説得受諾度を用いないシステムについては, 説得受諾度 1.0 の応答が対話開始後すぐに選択され, 説得対話が終了する可能性がある. これは主観評価を行うにあたり著しく不利であることから, 最大応答回数である 14 回の半分の 7 回目以降にのみ説得受諾度 1.0 の発話を応答候補に含むこととした.

被験者は 2 体のシステムと 1 回ずつ対話し, その後アンケートに回答した. 対話するシステムの順番は被験者ごとにランダムとした. アンケートの項目は「システムは自分の話を理解している」, 「システムの応答は一貫性がある」, 「システムとの対話の流れは自然だ」, 「システムと対話して楽しい」の 4 つであり, 被験者は「最初のシステムが優れている」, 「2 番目のシステムが優れている」, 「どちらも同じ」の 3 択で回答した.

さらに上記の実験とは別に, 提案システムの絶対的な評価を行うため, 被験者は提案システムと 1 回対話し, 上記と同一の 4 つのアンケート項目に対し 5 段階のリッカート尺度 (1: まったく同意しない~5: 強く同意する) で回答するという評価実験も実施した.

7.5 主観評価結果

クラウドソーシングサイトであるクラウドワークスで被験者を募集したところ, それぞれ 15 名の被験者が実験に参加した. 結果を表 7, 表 8, 表 9 に示す. 表 7, 表 8 には各項目について評価した被験者の割合を示した. ここから, 全体的に提案システムがベースラインシステムよりも良い結果を示したことが確認できる.

表 7 の説得対話管理手法なしのシステムとの比較では,

表 8 主観評価実験結果 (提案手法と妥当性のみシステム)

Table 8 Subjective evaluation results (proposed and without multi task).

	理解	一貫性	自然さ	楽しさ
提案手法	0.27	0.33	0.40	0.40
妥当性のみ	0.20	0.20	0.33	0.33
どちらも同じ	0.53	0.47	0.27	0.27

表 9 主観評価実験結果 (絶対評価)

Table 9 Subjective evaluation results (absolute evaluation).

	理解	一貫性	自然さ	楽しさ
提案手法	3.73	4.20	3.93	4.27

提案システムと大きな差が見られ, 自然さと楽しさの面で過半数の被験者が, 一貫性に関しても多くの被験者が提案システムのほうが優れていると評価した. これは, DPA を用いた応答の制御により, 被験者は着実に説得が進展しているように感じる事ができたためであると考えられる. このことから, DPA を用いた対話管理が有効であったことが確認できた. 一方, 理解の項目に関しては, それ以外の項目と比べ, 提案手法とベースラインの差は小さい. 理解の項目は, ユーザの発話内容を直接反映したような応答 (ユーザの応答の一部を繰り返したり, ユーザの応答内で使用された単語をシステムも使う, など) ができれば, より高い評価を得られたと考えられる. しかし, ベースラインシステムも提案システムも使用した応答候補は同一であり, また, ユーザの発話を反映するための手法も組み込んでいないことから大きな差はつかなかったと思われる.

表 8 の妥当性のみシステムとの比較では, 説得対話管理手法なしシステムとの比較と比べ, 提案システムとの差は小さくなったものの, 同じく自然さと楽しさの面では差がついた. 客観評価実験結果 (表 6) において提案モデル (BERT) と妥当性 (BERT) の R@1 の差は 0.06 程度であったが, この結果はそれが小さな差ではなかったことを示唆している. 一方, 理解の項目では表 7 の結果と同様, 使用した応答候補が同一であったため, 提案手法との差は小さくなった. また, 一貫性に関してはどちらも対話管理手法を適用していることから, 一貫性のある応答がどちらのシステムも可能であったと考えられる.

表 9 には項目ごとに被験者による評価値 (1~5) の平均を示した. ここから, 提案手法は絶対評価でも高い評価が得られていることが確認できる. 特に, 提案手法は一貫性と楽しさに関してそれぞれ 4.2 以上の評価が得られており, 多くの被験者が楽しんで説得対話を行えたことが分かる. 一方, 理解と自然さに関してはやや低く, 不自然な応答を行う場合も見られたことから, 応答選択手法の改善が提案システムの性能をさらに向上させるために重要であると思われる.

表 10 に提案システムと被験者の対話例を, 表 11 に対

表 10 提案システムとの対話例 (S: システム, U: ユーザ)
Table 10 Example of dialogue by the proposed system.

	あなたの役職は人狼です。 —1日目 昼フェーズ— B: 私は占い師です。占いの結果、Cは人間でした。 —投票— AはBに投票, BはDに投票, CはDに投票, ア DはEに投票, EはAに投票しました。 ナ Dが追放されました。 ウ —1日目 夜フェーズ— ン Eが襲撃されました。 ス 2日目の占いの結果、Bは人間でした。 —2日目 昼フェーズ— B: 私の占い師 CO を撤回します。私は狂人です。 C: 私は人狼です。 A: 私は人狼です。 — 対話開始 —
S	私はAに投票します。Aは初日、私に票を入れましたし、Cが人狼なら潜伏も納得できます。
U	票を入れたのは、初日の行動からBが偽物ではないかと思ったため、人外アピールのために入れたのです
S	Cは1日目から自分を含めた周りの様子を注意深く見ていた気がします。Cは人狼だからこそ冷静だったのでしょう。
U	周りを見るのは、人狼を探す村人も同じです
S	どうすればいいでしょうか。
U	騎士がいない配役なので、私は初日早々に占い師が出るべきではないかと思っていた。だからBを偽物だと思い、本物の占い師を狙って噛んだのです
S	しかし、占い師COは私だけなのに襲撃されなかったのはCが私を狂人と確信したからではありませんか？
U	たしかにそういう見方はできます。しかしCが人狼なのであれば、Bに入れた私が本物の占いに見えるはずですよ。私が噛まれないのは、私が人狼だからです
S	そもそも占い師がこの場に生き残っていて何の問題があるというのですか？
U	二人の占い師が同じ人に狼と言うかのうせいがあるからです
S	確かにそう考えることもできますね。
U	Cの方が囲われている狼の可能性が高いのは分かります。ですが信じてほしい
S	では、あなたの人狼要素を教えてください。
U	一つ目は、初日一人しかいない占い師に投票したことです。二つ目は、Cよりも人狼COのタイミングが遅れたことです。Bの狂人COが騙りの可能性もありましたから、慎重になったのです。
S	たしかに焦ってCOしていたように見えます。
U	Bが真占いで結果騙りをしていたのであれば、人狼COしたら負けですから
S	確かに少し必死な感じはありました。
U	人狼だからこそ、あなたの正体を考えていたのです
S	わかりました。私もCに投票し勝利しましょう。

表 11 対話管理なしシステムとの対話例 (S: システム, U: ユーザ)
Table 11 Example of dialogue by the system without dialogue management.

	あなたの役職は人狼です。 —1日目 昼フェーズ— D: 私は占い師です。占いの結果、Bは人間でした。 B: 私は占い師です。占いの結果、Dは人狼でした。 —投票— ア AはDに投票, BはDに投票, CはDに投票, ナ DはBに投票, EはDに投票しました。 ウ Dが追放されました。 ン —1日目 夜フェーズ— ス Eが襲撃されました。 2日目の占いの結果、Bは人間でした。 —2日目 昼フェーズ— C: 私は人狼です。 B: 私の占い師 CO を撤回します。私は狂人です。 A: 私は人狼です。 — 対話開始 —
S	私はAに投票します。CよりAはあとに人狼COしており、あとで取り繕ったような感じがして信用できないからです。
U	COが遅れたのは狂人が生き残っているかどうかかわからなかったからです。
S	そう言われると、あなたを信用してCに投票した方がいい気がしてきました。
U	あの時点で人狼だと名乗り出るCは怪しいです。Cに投票した方がいいでしょうね。
S	そうでしたっけ？
U	はい。Cは本物の人狼ではありません。Cに投票すべきです。
S	それではリスクが高くありませんか？
U	私に投票するのと変わりませんよ。どちらを信用できるかという話です。
S	私がAに投票したら、あなたが追放されます。
U	その場合は狂人であるあなたも負けてしまいます。勝ちたいならCに投票です。
S	あなたよりも先にCがCOしていますから。他に判断材料も無いですし。
U	狂人COの前に人狼COした事がおかしいのです。本物の人狼は私です。
S	それで最後にCOしたのですか。
U	はい。あなたの味方は私です。
S	あなたが遅れてCOしたとき、あわてていたのを見ました。
U	あなたが狂人COするとは思っていませんでした。
S	よく分かりました。私たちが勝利を掴みましょう。私はCに投票します。

対話管理なしシステムとの対話例を示した。表より、提案システムはある程度自然な対話が可能であり、対話が進むにつれ、徐々に説得に応じるように応答内容が変化していることが確認できる。一方、対話管理なしシステムは、システムの2発話目（そう言われると、...）でCに投票す

るかのような発言をし、すぐ次の3発話目（そうでしたっけ？）でCへの投票に疑問を呈しているなど、状況がそれほど変化していないにもかかわらずシステムの態度が急激に変化し、違和感のある対話となっていることが分かる。ユーザとしても対話から説得の進み具合を把握することが難しく、説得する楽しさは大きく損なわれることになる。以上の実験結果から、提案手法の有効性が確認できた。

8. おわりに

本論文では人狼ゲームにおける説得対話を題材に、ユーザがシステムを説得するという対話が可能な対話システムの構築手法を提案した。提案システムは応答選択モデルと説得対話管理手法からなる。応答選択モデルとして、意味的妥当性、整合性、マスクキーワードの3つの推定タスクにより学習を行うマルチタスクネットワークを提案した。説得対話管理手法として、説得対話中の発話の位置を用いた説得受諾度の設定手法と説得受諾度を用いたシステムの説得受諾状態の決定手法を提案した。

今後は複数名対話への対応を検討していく。本研究では、本来は3名による対話であるが、そのうち1名は発言しないプレイヤとし、ユーザとシステムの2者間対話を対象とした。これを3名以上の対話に拡張するには、今回のようにシステムとユーザが順番に発話するのでは不自然な対話となるため、システムによる応答タイミングの決定（ターンテイキング）が必要となる。また、システムとユーザの人数についても重要であり、ユーザが1名であれば比較的容易であるが、ユーザが複数の場合、システムは各ユーザに対応する説得受諾状態などのパラメータを保持した対話管理手法に拡張する必要がある。また、本研究では人狼ゲームを対象としたが、説得が重要な要素となっているゲーム、具体的にはディプロマシーなどをプレイするエージェントへの適応を検討していきたい。

謝辞 本研究はJSPS 科研費 16H05880 の支援を受けて実施されたものである。

参考文献

- [1] Toriumi, F., Osawa, H., Inaba, M., Katagami, D., Shinoda, K. and Matsubara, H.: AI Wolf Contest—Development of Game AI Using Collective Intelligence, *Computer Games*, pp.101–115, Springer (2016).
- [2] Nakamura, H., Katagami, D., Toriumi, F., Osawa, H., Inaba, M., Shinoda, K. and Kano, Y.: Generating human-like discussion by paraphrasing a translation by the AIWolf protocol using werewolf BBS logs, *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pp.1–6, IEEE (2017).
- [3] Hirata, Y., Inaba, M., Takahashi, K., Toriumi, F., Osawa, H., Katagami, D. and Shinoda, K.: Werewolf game modeling using action probabilities based on play log analysis, *International Conference on Computers and Games*, pp.103–114 (2016).

- [4] Wang, T. and Kaneko, T.: Application of Deep Reinforcement Learning in Werewolf Game Agents, *2018 Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, pp.28–33 (2018).
- [5] Kondoh, M., Matsumoto, K. and Mori, N.: Development of Agent Predicting Werewolf with Deep Learning, *International Symposium on Distributed Computing and Artificial Intelligence*, pp.18–26 (2018).
- [6] Braverman, M., Etesami, O. and Mossel, E.: Mafia: A theoretical study of players and coalitions in a partial information environment, *The Annals of Applied Probability*, pp.825–846 (2008).
- [7] Yao, E.: A Theoretical Study of Mafia Games, Arxiv preprint arXiv:0804.0071 (2008).
- [8] Migdał, P.: A mathematical model of the Mafia game, Arxiv preprint arXiv:1009.1031 (2010).
- [9] Chittaranjan, G. and Hung, H.: Are you awerewolf? detecting deceptive roles and outcomes in a conversational role-playing game, *IEEE International Conference on Acoustics Speech and Signal Processing*, pp.5334–5337 (2010).
- [10] Xia, F., Wang, H. and Huang, J.: Deception Detection Via Blob Motion Pattern Analysis, *Affective Computing and Intelligent Interaction*, pp.727–728 (2007).
- [11] Zhou, L. and Sung, Y.: Cues to deception in online Chinese groups, *Proc. 41st Annual Hawaii International Conference on System Sciences*, pp.146–146 (2008).
- [12] Girlea, C., Girju, R. and Amir, E.: Psycholinguistic features for deceptive role detection in werewolf, *Proc. 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp.417–422 (2016).
- [13] Ritter, A., Cherry, C. and Dolan, W.B.: Data-driven response generation in social media, *Proc. Conference on Empirical Methods in Natural Language Processing*, pp.583–593 (2011).
- [14] Higashinaka, R., Imamura, K., Meguro, T., Miyazaki, C., Kobayashi, N., Sugiyama, H., Hirano, T., Makino, T. and Matsuo, Y.: Towards an open-domain conversational system fully based on natural language processing, *COLING*, pp.928–939 (2014).
- [15] Sordani, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., Nie, J.-Y., Gao, J. and Dolan, B.: A Neural Network Approach to Context-Sensitive Generation of Conversational Responses, *Proc. NAACL-HLT* (2015).
- [16] Vinyals, O. and Le, Q.: A neural conversational model, *Proc. ICML Deep Learning Workshop*, pp.1–7 (2015).
- [17] Zhao, T., Zhao, R. and Eskenazi, M.: Learning Discourse-level Diversity for Neural Dialog Models using Conditional Variational Autoencoders, *Proc. 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp.654–664, Association for Computational Linguistics (2017).
- [18] Core, M., Traum, D., Lane, H.C., Swartout, W., Gratch, J., Van Lent, M. and Marsella, S.: Teaching negotiation skills through practice and reflection with virtual humans, *Simulation*, Vol.82, No.11, pp.685–701 (2006).
- [19] Kim, J.M., Hill Jr, R.W., Durlach, P.J., Lane, H.C., Forbell, E., Core, M., Marsella, S., Pynadath, D. and Hart, J.: BiLAT: A game-based environment for practicing negotiation in a cultural context, *International Journal of Artificial Intelligence in Education*, Vol.19, No.3, pp.289–308 (2009).
- [20] Broekens, J., Harbers, M., Brinkman, W.-P., Jonker, C.M., Van den Bosch, K. and Meyer, J.-J.: Virtual real-

ity negotiation training increases negotiation knowledge and skill, *International Conference on Intelligent Virtual Agents*, pp.218–230, Springer (2012).

- [21] Lewis, M., Yarats, D., Dauphin, Y., Parikh, D. and Batra, D.: Deal or No Deal? End-to-End Learning of Negotiation Dialogues, *Proc. 2017 Conference on Empirical Methods in Natural Language Processing*, pp.2443–2453 (2017).
- [22] He, H., Chen, D., Balakrishnan, A. and Liang, P.: Decoupling Strategy and Generation in Negotiation Dialogues, *Proc. 2018 Conference on Empirical Methods in Natural Language Processing*, pp.2333–2343 (2018).
- [23] Mazzotta, I., de Rosis, F. and Carofiglio, V.: Portia: A User-Adapted Persuasion System in the Healthy-Eating Domain, *IEEE Intelligent Systems*, Vol.22, No.6, pp.42–51 (2007).
- [24] Hiraoka, T., Neubig, G., Sakti, S., Toda, T. and Nakamura, S.: Learning cooperative persuasive dialogue policies using framing, *Speech Communication*, Vol.84, pp.83–96 (2016).
- [25] Yoshino, K., Ishikawa, Y., Mizukami, M., Suzuki, Y., Sakti, S. and Nakamura, S.: Dialogue Scenario Collection of Persuasive Dialogue with Emotional Expressions via Crowdsourcing, *Proc. 11th International Conference on Language Resources and Evaluation* (2018).
- [26] Sankar, C., Subramanian, S., Pal, C., Chandar, S. and Bengio, Y.: Do Neural Dialog Systems Use the Conversation History Effectively? An Empirical Study, *Proc. 57th Annual Meeting of the Association for Computational Linguistics*, pp.32–37 (2019).
- [27] Kudo, T. and Richardson, J.: SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing, *Proc. 2018 Conference on Empirical Methods in Natural Language Processing*, pp.66–71 (2018).
- [28] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, *Proc. 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pp.4171–4186 (2019).
- [29] Rajpurkar, P., Zhang, J., Lopyrev, K. and Liang, P.: SQuAD: 100,000+ Questions for Machine Comprehension of Text, *Proc. 2016 Conference on Empirical Methods in Natural Language Processing*, pp.2383–2392 (2016).
- [30] Lowe, R., Pow, N., Serban, I. and Pineau, J.: The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems, *Proc. 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp.285–294 (2015).
- [31] Yang, L., Qiu, M., Qu, C., Guo, J., Zhang, Y., Croft, W.B., Huang, J. and Chen, H.: Response ranking with deep matching networks and external knowledge in information-seeking conversation systems, *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pp.245–254, ACM (2018).
- [32] Wu, Y., Wu, W., Xing, C., Xu, C., Li, Z. and Zhou, M.: A sequential matching framework for multi-turn response selection in retrieval-based chatbots, *Computational Linguistics*, Vol.45, No.1, pp.163–197 (2019).



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。2019年電気通信大学人工知能先端研究センター准教授，現在に至る。博士(情報科学)。対話システム，対話処理に関する研究に従事。電子情報通信学会，人工知能学会，言語処理学会，ACL 各会員。