

# 人狼ログからの小説の自動生成

松山 諒平 佐藤 理史 松崎 拓也

名古屋大学大学院 工学研究科 電子情報システム専攻

## 1 はじめに

2012年に、ショートショートを対象とした文学賞として日経「星新一賞」が新設された。この賞は、応募要領に、人間以外(人工知能等)からの応募作品も受け付けると明記されている点に特徴がある[1]。2015年9月末に締め切られた第3回「星新一賞」には、コンピュータを利用して作成された小説11編が応募され、現在までに、そのうちの4作品がウェブ上で公開されている[2, 3]。

小説を作成するためには、プロットを作成し、それを文章化することが必要である。4作品のうちの2作品を作成した「人狼知能プロジェクト」は、プロット作成の部分を自動化し、得られたプロットを人間が文章化した。残りの2作品を作成した「きまぐれ人工知能プロジェクト 作家ですよ」は、人間がプロットを作り、それをコンピュータを用いて文章化した。

我々は、2016年9月末締切の第4回「星新一賞」応募にあたり、2つのプロジェクトの成果を統合することを試みた。すなわち、コンピュータによって作成されたプロットを、コンピュータによって文章化するシステムを作成した。本論文では、このシステムについて述べる。

## 2 システムの入出力

### 2.1 システムの入力

システムの入力は、人狼知能プロジェクトから提供を受けた、複数のエージェント(コンピュータプログラム)による人狼ゲームの記録である[4]。以下では、これを人狼ログまたは、単にログと呼ぶ。図1にログの一部を示す。本研究では、このようなログを小説のプロットとして利用する。

人狼ログは、ゲームマスターの視点から見たゲームの記録である。それぞれの行は、あるエージェントの一つの行動の記録であり、以下の要素から構成される。

1. その行動が行われた日
2. 行動タイプ
3. 行動の詳細情報

たとえば、図1の先頭の行は、次のことを表す。

```

1,talk,5,10,COMINGOUT Agent[10] MEDIUM
1,talk,6,3,DIVINED Agent[01] HUMAN
1,talk,7,8,VOTE Agent[10]
1,talk,8,5,VOTE Agent[02]
1,talk,9,4,VOTE Agent[10]
1,whisper,0,7,VOTE Agent[10]
1,whisper,1,6,Over
1,divine,3,9,HUMAN
1,guard,1,3,SEER
1,attackVote,6,1
1,attackVote,7,3
1,execute,2,VILLAGER
1,attack,3,false
    
```

図1: 人狼ログの例

- 1: 1日目に行われた
- talk: 行動タイプは発話行動
- 5: この発話は、その日の5番目の発話
- 10: 発話者(人狼エージェント)のID
- COMINGOUT Agent[10] MEDIUM: 発話内容  
(エージェント10(自分)は霊媒師である)

なお、人狼ログのサイズは、多くのものが500行から800行である。ゲーム参加者(エージェント)は10名で、人狼2名、霊媒師、占い師、狩人、狂人がそれぞれ1名、残りの4名が村人である。

提供されたログのうち、システムの入力として採用するものは「最後の会議の参加者が3名であるログ」に限定する。この理由については後述する。

### 2.2 システムの出力

システムの出力は、以下に示す四つのパートから構成される短編小説である。

1. 導入: 背景と人狼ゲームの説明
2. ゲーム進行描写1: あるエージェント(人狼)から見たゲームの進行
3. ゲーム進行描写2: 別のエージェント(村人側)から見たゲームの進行
4. 結末: ゲームの結果とエピローグ

これらのパートのうち、二つのゲーム進行描写は、それぞれのエージェントの視点(つまり、一人称)で語られる。実際に、システムが出力した作品の一部(ゲーム進行描写1)を図2に示す。

ゲーム進行描写は、次の二つの部分で構成する。

私はF恵。普段は将棋の棋士 AI として働いている。私はあらゆる手を想定でき、未だに負けたことが無い。今日、私は人狼テストを受けさせられることになった。このような遊びをして何の意味があるのだろうか。開発者はやる気だが、私はあまり乗り気ではなかった。

プレイヤーは全部で10人。A太、B香、C子、D亮、E美、G郎、H夫、I子、J也、それと私F恵だ。有名なAIばかり。同じAIとして負けたくない。

テストが始まった。私の役職は「人狼」だった。下手なことをしないように気をつけよう。

もう一人の人狼はG郎だった。私はG郎をちらりと見て、さりげなくお互いを確認した。

1日目。C子が占い師だと名乗り出た。つづいて、J也が霊媒師だと名乗り出た。C子は占いの結果、A太が人間だったと伝えた。誰を処刑するべきか意見は様々あったが、投票の結果、B香が処刑されることになった。私はC子を襲撃することにした。しかし、襲撃は狩人に防がれてしまった。

2日目。D亮が霊媒師だと名乗り出た。C子は占いの結果、I子が人間だったと伝えた。J也は霊媒の結果、B香が人間だったと伝えた。D亮は霊媒の結果、B香が人間だったと伝えた。会議の中で霊媒師を疑う声上がり、J也が処刑された。私はI子を襲撃することにし、襲撃に成功した。

3日目。C子は占いの結果、G郎が人狼だったと伝えた。D亮は霊媒の結果、J也が人狼だったと伝えた。会議の中で霊媒師に嘘付きがいるという声上がり、D亮が処刑された。私はA太を襲撃することにし、襲撃に成功した。

4日目。C子は占いの結果、E美が人間だったと伝えた。投票の結果、仲間のG郎が処刑されてしまった。これで残る人狼は私だけとなった。私はE美を襲撃することにし、襲撃に成功した。

5日目が始まった。ここまでゲームが進行し、生き残っているのは、C子、私、H夫のたった3人だけとなった。今日の処刑で勝敗が決するというのは誰の目にも明らかだ。互いに見合っている目には緊張の色がうかがえた。

残りは三人。この最後の処刑を生き残れば私の勝利だ。一番の懸念事項は占い師の可能性の高いC子が生き残っていることだ。

会議の始めにC子が占いの結果について話した。

「占いの結果、F恵は人狼だった」  
この最終局面で人狼であることを占い師に看破されてしまった。しかし、占い師が生き残ってしまった以上それはしょうがないことだ。私はまだあきらめていない。

「人狼は、F恵だ。F恵を処刑したら人間の勝利だ」  
C子は言った。いやまだ、勝利を宣言するには早いだろう。私は反論した。

「C子こそ人狼だ。占い師だったらとっくに人狼に襲撃されているはずだ。本当は人狼だからこそ襲撃もされずに生き延びているのではないか」

投票は多数決。せつかくの占い師も、H夫から信用が得られなければ、勝利することはできない。さらに、H夫にはC子が間違いなく占い師であると判断できるだけの証拠はないだろう。C子へ疑いを持たせることさえできれば、C子を処刑できる可能性もある。

それからしばらく沈黙の時間が流れた。その間にいろいろな可能性が頭の中を巡る。もしかしらC子は狂人かもしれないという考えも思いつくが、ここはおとなしくすることにした。そのまま会議の時間は終了し、投票となった。H夫はどちらに入れるだろうか。私は祈るように結果を待った。

図 2: ゲーム進行描写 1 の例

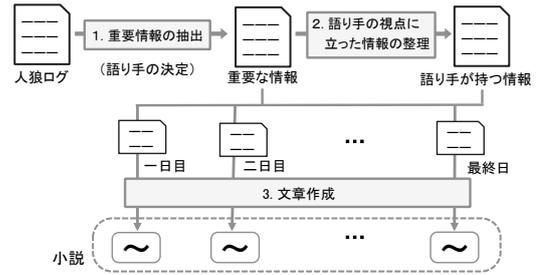


図 3: システムの全体像

### 1. 最後の会議までの経緯：

それぞれの日に起きた重要な出来事を記述し、ゲームがどのように進行したかを簡潔に描写する。

### 2. 最後の会議の描写：

エージェント同士の会話とエージェントの心情を記述し、ゲームの様子を詳細に描写する。

このような構成を採用した理由は、ゲームの決着を左右する「最後の会議」がゲームのハイライトであり、これを物語の山場とするためである。

## 3 システムの構成

### 3.1 システムの概要

人狼ログから小説を作成するために、システムは次の三つのことを行う必要がある。

#### 1. 重要情報の抽出：

ログには、ゲーム中のすべての出来事が記述されているため、それらを逐一文章化しても小説にはならない。そのため、小説の素材となる重要な情報のみを抽出することが必要である。

#### 2. 不足情報の追加：

ログには、エージェントがどのような性格か、ゲーム中何を考えていたのか、等の情報は一切記述されていない。そのため、必要に応じてこれらの情報を追加することが必要である。

#### 3. 語り手の視点に立った情報の整理：

ログは、ゲームマスターの視点で記述されている。ゲーム進行描写をエージェント視点で書くためには、抽出した情報を文章の語り手の視点に立って整理する必要がある。

システムの全体像を図 3 に示す。システムは、(1) 重要情報の抽出、(2) 語り手の視点に立った情報の整理、(3) 文章作成、の 3 つのモジュールで構成される。文章生成では、文章生成器 GhostWriter[5] を使用し、不足情報を追加しつつ文章を作成する。

### 3.2 重要情報の抽出

ログの行動タイプに基づき、重要な行動のみを抽出する。行動タイプ (全 11 種類) の重要度は表 1 に示す

表 1: 行動タイプの重要度

重要度	行動タイプ	意味
	divine	占い先指定
	guard	護衛先指定
	execute	処刑結果
	attack	襲撃結果
	result	ゲームの結果
	status	エージェントの状態
	talk	発話
×	whisper	人狼のみの会話
×	vote	処刑投票
×	attackvote	襲撃先指定

DAY: 5 語り手: Agent6

Agent1:	情報なし
Agent2:	カミングアウト(村人)
Agent3:	カミングアウト(占い師)
Agent4:	カミングアウト(霊媒師)
Agent5:	情報なし
Agent6:	自分
Agent7:	仲間の人狼
Agent8:	情報なし
Agent9:	情報なし
Agent10:	カミングアウト(霊媒師)

図 4: 整理した「語り手が持つ情報」の例

表 2: 発話内容の重要度

重要度	発話内容	意味
	DIVINED	占いの結果
	INQUESTED	霊媒の結果
	GUARDED	護衛した
	COMINGOUT	自分の役職を表明
×	ESTIMATE	他人の役職の予想
×	VOTE	投票予定
×	AGREE	発話への同意
×	DISAGREE	発話への反対
×	Skip	まだ話すことがある
×	Over	話すことは無い

ように定めた。表の記号は以下の意味を持つ。

: すべて抽出する

: 一部抽出する

×: 抽出しない

一部抽出する行動は、その詳細情報に基づいて抽出するか否かを決定する。行動タイプ “status” は、エージェントの状態が “ALIVE” の場合のみ抽出する。行動タイプ “talk” は、表 2 に基づいて抽出するか否かを決定する。

### 3.3 語り手の視点に立った情報の整理

ある時点(日付)までに、語り手となるエージェントが得た情報を、ゲーム参加エージェントごとに整理する。例を図 4 に示す。この例は、ゲーム 5 日目までにエージェント 6 が得た情報を整理したものである。

整理する情報は、語り手となるエージェントが持つ「他のエージェントが敵か味方が判断するための情報」である。具体的には以下の二種類の情報がある。

#### 1. 役職のカミングアウト:

全体で占い師が一人しか存在しない状況で、占い師だと名乗るエージェントが複数いる場合、その中に人狼陣営のエージェントがいる可能性が高い。

#### 2. 特定の役職のエージェントのみ知る情報:

特定の役職は、他のエージェントの正体を知り得る。たとえば、占い師は、占いにより特定のエージェントが人狼かどうか知ることができる。人狼は、自分以外の人狼が誰かを知ることができる。

これらの情報は、人狼ゲームの戦略上重要な、他のエージェントが敵か味方が判断する手がかりであり、エージェントの行動を大きく左右する。

## 3.4 文章の作成

### 3.4.1 最後の会議までの経緯

この部分は以下の内容で構成する。

1. 語り手のエージェントの紹介
2. 他のエージェントの紹介
3. ゲーム開始、役職の確認
4. 他のエージェントの様子を確認
5. 1 日目から 4 日目のあらすじ

はじめの四つの内容の描写では、ログから得た情報のうち、語り手のエージェントの ID と役職のみを使用する。他の情報(たとえば、エージェントの名前が F 恵であること)は、事前に用意した設定から追加する。

1 日目から 4 日目のあらすじは、一日ずつ作成する。一日分のあらすじの文章構造は四日とも同じである。その文章構造を図 5 に示す。この図に示すように、「会議」を構成する内容は三つある。このうち、「カミングアウトについて」と「占い/霊媒結果について」は、それぞれ、カミングアウトまたは占い/霊媒結果の発言があった日のみ記述する。それらの発言がすべてなかった日は、「新しい情報はなかった」と記述する。

それぞれの内容に対しては、その内容を記述するためのテンプレートが用意されており、ログから得られた情報に沿って、テンプレートの選択と穴埋めを行うことで文章を作成する。例として、「襲撃について」を記述するためのテンプレートの一部を表 3 に示す。この例では、語り手の役職と襲撃の成否に基づいてテンプレートを選択し、襲撃された人物の名前を空欄に埋め込む。

### 3.4.2 最後の会議の描写

「最後の会議までの経緯」では、1 文から 3 文程度の短いテンプレートを用いて文章を作成するのに対し、「最後の会議の描写」では、一つの大きなテンプレ

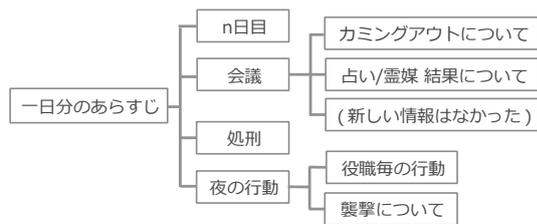


図 5: 一日分あらすじの文章構造

表 3: テンプレートの例

語り手	襲撃結果	テンプレート
村人	成功	今晚襲撃されたのは〇〇だった
村人	失敗	今晚襲撃された人はいなかった
人狼	-	私は〇〇を襲撃することにした

トで文章を作成する。このような方法を採用した理由は、「最後の会議の描写」の文章構造が、会議の流れによって定まるからである。

想定される会議の流れとして、以下のような例がある。

- 人狼を特定する明確な手がかりはない。生き残った3名は、自分は人狼ではないと主張する。
- 3名のうち、人狼と疑わしい2名が、互いにもう一方が人狼だと主張する。残りの1名はどちらの主張を信じるか悩む。

これらのような会議の流れが16種類設定されており、それらはさらに、会議中での語り手の立場などによって細分化されている。テンプレートは合計で86種類ある。テンプレートの選択には、以下の三つの情報を主に使用する。

1. 会議に参加するエージェントのカミングアウト：各エージェントがどの役職のフリをしているかという情報。会議の流れを決定する。
2. 会議の結果、処刑されるエージェント：会議が誰を処刑する方向に進むかを決定する。
3. 語り手のエージェントの役職：会議中での語り手の立場を決定する。

現在は、会議参加者の数が固定されたテンプレートを用いているため、テンプレートの選択時に参加者の数が定まっている必要がある。提供されたログ198個のうち、158個のログでは最後の会議の参加者が3名であり、残りのログでは4名であった。そこで、今回は、「最後の会議の参加者が3名であるログ」のみを入力として採用することとした。

## 4 おわりに

作成したシステムは、入力として採用した158個の人狼ログのうち、153個に対して小説を生成すること

ができる。残りの5個に対しては、「最後の会議の描写」を、用意したテンプレートでは生成できない。これらのログでは、占い師だと名乗る人狼が最後の会議まで生存する。このような状況に対しても、新たなテンプレートを追加すれば、「最後の会議の描写」が可能となる。

より重大な課題は、生成された小説には、次のような要素が不足しているという点である。

- キャラクターの設定：  
F 恵は普段は棋士として働いている、などの簡単な設定はあるが、考え方や話し方にまったく反映されていない。そのため、個性に欠ける。
- 小説のテーマ：  
一般に、小説には、成長、友情、恋愛などのテーマが必要である。しかし、現在の作成手法は、テーマについてまったく考慮していない。

本研究では、人狼ログを素材として小説を生成するが、ログから得られる情報だけでは不十分である。読むに耐える小説を生成するためには、上記のような要素を十分に組み込むことが必要である。

## 謝辞

人狼ログを提供していただいた人狼知能プロジェクトに深く感謝します。本研究は、JSPS 科学研究費基盤研究(B)「文章の読解と産出のための言語処理技術」(課題番号 15H02748)の助成を受けている。

## 参考文献

- [1] 第4回 日経「星新一賞」. <http://hoshiaward.nikkei.co.jp/index.html>.
- [2] 人狼知能プロジェクト 第3回日経星新一賞応募作公開. <http://aiwolf.org/archives/873>.
- [3] きまぐれ人工知能プロジェクト作家ですよ > 成果. [http://www.fun.ac.jp/~kimagure\\_ai/results/](http://www.fun.ac.jp/~kimagure_ai/results/).
- [4] 鳥海不二夫, 片山大輔, 大澤博隆, 稲葉通将, 篠田孝祐, 狩野芳伸. 人狼知能 だます・見破る・説得する人工知能. 森北出版, 2016.
- [5] 佐藤理史. コンピュータが小説を書く日. 日本経済新聞出版社, 2016.