

# 既存小説に依存しない小説の自動生成に関する一考察

福田 清人      森 直樹      松本 啓之亮

大阪府立大学 工学研究科

{fukuda@ss., mori@, matsu@}cs.osakafu-u.ac.jp

## 1 はじめに

近年、絵画や音楽、小説といった人間の感性に基づく創作物を計算機で自動生成することが人工知能の分野で重要な課題となっている。その中で、星新一らしい小説を自動生成することを目指した「気まぐれ人工知能プロジェクト『作家ですよ』」[1] が 2012 年にスタートするなど、小説の自動生成は自然言語処理や人工知能の多岐にわたる技術が必要とするチャレンジングな問題として注目を集めており、今後その重要性はますます高まると思われる。

現在、小説の自動生成に関する研究は既存の小説のストーリーを利用した研究が主流となっており、既存の小説から抽出したストーリーとなる文章を一部変更することで小説を自動生成する研究 [2] や、プロットの文学理論を利用した物語の自動生成システムの研究 [3] が存在する。しかしながら、これらの研究では利用した小説のストーリーに大きく依存してしまっていることが問題点として挙げられる。

以上の点より、既存の小説を利用しない小説の自動生成を最終目的とする。その前段階として、Agent-Based simulation (ABS) により生成したログデータを用いてストーリーを半自動生成する手法を提案し、生成されたストーリーから人手で生成した小説を解析することで既存の小説を利用しない小説の自動生成に必要な要素について考察する。

以下に本研究の構成を示す。第 2 章で提案手法について説明し、第 3 章で動作実験について結果と考察を示す。最後に第 4 章でまとめと今後の課題について述べる。

## 2 提案手法

本研究では計算機が創発的にストーリーを生成することを目標として、計算機で得られるログデータを用いたストーリーの半自動生成手法を提案する。ここで、各登場人物に関する一連の状態変化を多様な文で言語

化したものを小説と定義し、小説のうち、状態変化を行動で言語化したものをストーリーと定義する。

### 2.1 提案手法の流れ

以下に提案手法のアルゴリズムを示す。それぞれ 1. ~ 3. については 2.2 ~ 2.4 に詳述する。

1. ログデータ生成 Agent-Based Simulation (ABS) によってログデータを生成する。
2. イベント生成 得られたログデータからストーリーを構成するイベントを生成する。
3. ストーリーベクトル生成 得られたイベントからストーリーをベクトル形式で表したストーリーベクトルを生成する。
4. ストーリー生成 生成されたストーリーベクトルから人手でストーリーを生成する。

### 2.2 ログデータの生成

本研究では ABS によってログデータを生成する。ABS には提案手法のために実装したものをを用いる。

#### 2.2.1 ログデータの種類

ABS によってゲームログとステータスログという 2 種類のログデータを生成する。

**ゲームログ** ABS においてエージェントが特定の行動をするたびに、ターン、シーン、主体となるエージェント名、客体となるエージェント名、行動の種類、行動による友好度の変化量、行動による感情の変化量 という 7 種類のデータをゲームログとして取得する。

**ステータスログ** 各ターンの終了時にターンと全エージェントの全パラメータをステータスログとして取得する。

## 2.2.2 ABS のアルゴリズム

以下に実装した ABS のアルゴリズムについて示す。ここで、平均が  $\mu$ 、分散が  $\sigma^2$  の正規乱数を  $N(\mu, \sigma^2)$  とする。

1. フィールドサイズを  $x_{\max} \times y_{\max}$ 、最大ターン数を  $t_{\max}$ 、参加エージェント数を  $n$  とする。
2. 環境エージェントを生成する。環境エージェントは位置  $x, y$  に対してシーン番号  $s_{xy}$ 、環境値  $a_{xy}, b_{xy,t}$  をパラメータとして持つ。ここで、環境値  $a_{xy}$  は  $-1$  か  $1$  の二値であり、環境値  $b_{xy,t}$  は  $-5 \leq b_{xy,t} \leq 5$  を満たす実数である。また、環境値の初期値  $b_{xy,0}$  は  $-3 \leq b_{xy,0} \leq 3$  を満たす実数と定める。
3. 参加エージェントを  $n$  体生成する。各エージェントは名前、位置  $x_t, y_t$ 、性別  $m$ 、各エージェントに対する友好度  $F_t = \{f_{1,t}, f_{2,t}, \dots, f_{n,t}\}$ 、各エージェントに対する告白フラグ  $L = \{\ell_1, \ell_2, \dots, \ell_n\}$ 、感情  $E_t = \{e_{1,t}, e_{2,t}\}$  をパラメータとして持つ。友好度  $f \in F$  は  $0 \leq f \leq 100$  を、感情  $e \in E$  は  $-100 \leq e \leq 100$  をそれぞれ満たす実数であり、性別  $m$  および告白フラグ  $\ell \in L$  は  $0$  か  $1$  の二値である。また、それぞれ  $40 \leq f_0 \leq 60, -10 \leq e_0 \leq 10$  である。ここで、参加エージェント集合を  $P$  とする。
4. ターン  $t = 0$  とし、 $t$  における未行動のエージェント集合を  $P_t$ 、行動済みのエージェント集合を  $P'_t$  とする。
5.  $P_t = P, P'_t = \Phi$  とする。
6.  $P_t$  からランダムにエージェントを 1 体選択し、 $i$  とする。  $P_t = P_t \setminus i$  とし、  $P'_t = P'_t \cup i$  とする。
7. エージェント  $i$  をムーア近傍のいずれかに移動させる。エージェントはターンとシーン番号を要素とするスケジュールと目標位置を持ち、目標位置まで移動すると、スケジュールからターンにおけるシーン番号を取得し、そのシーンの中からランダムに目標位置を再設定する。ここで、シーンとはあるシーン番号を与えられた場所の集合というふうに定義する。
8. 自身以外のエージェントからランダムに 1 体選択し、 $j$  とする。エージェント  $i$  のエージェント  $j$  に対する友好度  $f_{j,t}^i$  および感情  $e_t^i \in E_t^i$  を以下の式に従って更新する。これは時間経過による心

理的な影響を表現するためである。

$$f_{j,t}^i = f_{j,t}^i - \frac{|N(0,1)| \times a_{xy}^i (f_{j,t}^i - 50)}{200} \quad (1)$$

$$e_t^i = e_t^i - \frac{|N(0,1)| \times a_{xy}^i e_t^i}{500} \quad (2)$$

9. エージェント  $i$  のノイマン近傍に他のエージェントが存在すれば、その中からランダムにエージェントを選択して  $k$  とする。存在しなければ 12 へ。
10. エージェント間の相互作用として、エージェント  $k$  のエージェント  $i$  に対する友好度  $f_{i,t}^k$  および感情  $e_{1,t}^k, e_{2,t}^k$  を以下の式に従って更新する。エージェント間の相互作用に対してゲームログを保存する。

$$f_{i,t+1}^k = f_{i,t}^k + \Delta f_{i,t}^k \quad (3)$$

$$e_{1,t+1}^k = e_{1,t}^k + \Delta e_{1,t}^k \quad (4)$$

$$e_{2,t+1}^k = e_{2,t}^k + \Delta e_{2,t}^k \quad (5)$$

$$\Delta f_{i,t}^k = |N(0,1)| \times a_{xy}^k d_{i,t}^k + b_{i,t}^k \quad (6)$$

$$d_{i,t}^k = \frac{2f_{i,t}^k + f_{k,t}^i}{20} + \frac{e_{1,t}^k (e_{2,t}^k + 100)}{1500} \quad (7)$$

$$\Delta e_{1,t}^k = |N(0,1.5)| \times \Delta f_{i,t}^k \quad (8)$$

$$\Delta e_{2,t}^k = N(0,1.5) \times |\Delta f_{i,t}^k| \quad (9)$$

11. 10 で得られた  $\Delta f_{i,t}^k$  または、エージェントを入れ替えた  $\Delta f_{k,t}^i$  が閾値  $h$  を超えていれば、エージェント  $i, k$  は自身以外のエージェントからランダムに 1 体選択し、そのエージェントに対する告白フラグを 1 ならば 0, 0 ならば 1 へと変化させる。ここで、 $h$  は閾値である
12.  $P_t \neq \Phi$  であれば、5 へ。
13.  $P'_t$  からランダムにエージェントを 1 体選択し、 $p$  とする。エージェント  $p$  のノイマン近傍に他のエージェントが存在すればランダムに 1 体選択し、 $p'$  とする。エージェント  $p$  のエージェント  $p'$  に対する告白フラグ  $\ell_{p'}^p$  が 1 であれば、 $\frac{|f_{p',t}^p - 50|}{100}$  の確率で  $p$  が  $p'$  に告白する。告白は以下の式で表される成功率  $r$  で成功する。告白すれば成功してもしなくてもゲームログを保存し 15 へ。告白しなければ  $P'_t = P'_t \setminus p$  とする。

$$r = \begin{cases} \frac{0.8f_{p',t}^p + 0.2\alpha}{200} & (m^p = m^{p'}) \\ \frac{0.8f_{p',t}^p + 0.2\alpha}{100} & (m^p \neq m^{p'}) \end{cases} \quad (10)$$

$$\alpha = \frac{e_{1,t}^{p'} \times e_{2,t}^{p'}}{100} \quad (11)$$

14.  $P'_t \neq \Phi$  であれば, 13 へ.
15. 環境エージェントの環境値  $b_{xy,t}$  を以下の式に従って更新する.

$$X = \frac{2(t+1)\pi}{500} \quad (12)$$

$$b_{xy,t+1} = b_{xy,0} + \sin(aX) + \cos(bX) \quad (13)$$

ここで,  $a, b$  はそれぞれ  $0 \leq a, b \leq 500$  を満たす定数である.

16.  $t = t + 1$  としてステータスログを保存する.  $t \geq t_{\max}$  ならば ABS を終了する. そうでなければ 5 へ.

## 2.3 イベントの取得

得られたゲームログをそのまま用いてストーリーベクトルを生成してしまうと, 同一の行動を連続して行ってしまう, 不自然なストーリーになってしまう. そこで, ゲームログに対して何らかの操作を適用してゲームログからストーリーを構成するイベントを取得する. 今回ゲームログを圧縮することでイベントを取得する手法を用いる. 以下に圧縮によるイベントの取得アルゴリズムを示す.

1. 得られたゲームログ集合を  $G$  とする.
2.  $G$  から, 連続した同一のシーンで取得されたゲームログを抽出し  $G'$  とする.
3.  $G'$  を客体となるエージェント名, 主体となるエージェント名, 行動の種類が同一のゲームログ集合に分割する.
4. 分割したゲームログ集合から 1 つずつゲームログを取得し, 前後のゲームログでターンの差が 50 未満であるすべてのゲームログから新たなゲームログを生成する. 生成されたゲームログにおいて, ターンおよびシーンはターンの値が小さいゲームログのものとし, 友好度の変化量および感情の変化量は生成に用いるすべてのゲームログの和とする.
5.  $G = G \setminus G'$  とし,  $G \neq \Phi$  であれば 2 へ.
6. で得られたゲームログをイベントとして取得する.

## 2.4 ストーリーベクトルの生成

取得したイベントを用いて直接ストーリーを生成することは困難であるため, イベントとステータスログ

からストーリーベクトル  $s$  を生成し, ストーリーベクトル  $s$  からストーリーを生成する. 以下にストーリーベクトルの生成アルゴリズムを示す.

1. 主体となるエージェントを  $i$ , 客体となるエージェントを  $j$  とし, イベントからターン  $t$ , 友好度の変化量  $\Delta f_t$ , 各感情の変化量  $\Delta e_t$  を取得する. また, ステータスログから  $t$  におけるエージェント  $i, j$  の, エージェント  $j, i$  に対する友好度  $f_{j,t}^i, f_{i,t}^j$  および感情  $e_t^i, e_t^j$  を取得する.
2. ストーリーベクトル  $s$  を以下のように定義する.

$$s = (s_{j,t}^i, s_{i,t}^j, s_{\Delta f_t}, s_{\Delta e_t}) \quad (14)$$

$$s_{j,t}^i = (v_{j,S}^i, v_{j,M}^i, v_{j,L}^i) \quad (15)$$

$$s_{i,t}^j = (v_{i,S}^j, v_{i,M}^j, v_{i,L}^j) \quad (16)$$

$$s_{\Delta f_t} = (v_{f,S}, v_{f,L}) \quad (17)$$

$$s_e = (v_{e1}, v_{e2}, v'_{e1}, v'_{e2}) \quad (18)$$

ここで, それぞれは以下のように求められる.

$$\begin{cases} v_{j,S}^i = 1 & (f_{j,t}^i \leq 30) \\ v_{j,M}^i = 1 & (30 < f_{j,t}^i < 70) \\ v_{j,L}^i = 1 & (f_{j,t}^i \geq 70) \end{cases} \quad (19)$$

$$\begin{cases} v_{f,S} = a & (0 < |\Delta f_t| < 20) \\ v_{f,L} = a & (|\Delta f_t| \geq 20) \end{cases} \quad (20)$$

$$a = \begin{cases} 1 & (\Delta f_t \geq 0) \\ -1 & (\Delta f_t < 0) \end{cases} \quad (21)$$

$$v_{ek} = \begin{cases} 2 & (e_{ki,t}^j \geq 5) \\ 1 & (e_{ki,t}^j \leq -5) \\ 0 & (-5 < e_{ki,t}^j < 5) \end{cases} \quad (22)$$

$$v'_{ek} = \begin{cases} 2 & (e_{ki,t}^j + \Delta e_{k,t} \geq 5) \\ 1 & (e_{ki,t}^j + \Delta e_{k,t} \leq -5) \\ 0 & (-5 < e_{ki,t}^j + \Delta e_{k,t} < 5) \end{cases} \quad (23)$$

## 3 動作実験

### 3.1 実験条件

実験条件は実験 1, 実験 2 とともに同一の実験条件を用いた. 短編小説の執筆経験がある 20 代の女性 2 名 (被験者 A, B) に対して実験した. 表 1 に実験条件を示す.

### 3.2 実験 1

提案手法を用いて生成されたストーリーベクトルからストーリーを生成してもらった. ストーリーの生成

ターン	自分	相手	場所	行動	行動番号	自分の友好度	相手の友好度	自分の感情変化	相手の感情変化	態
81	ヒロ	カナ	lectureB	友好度が下がる。(−8.34)	0	normal (40.27)	normal (38.56)	「」→「」	「怒」→「怒」	0
205	カナ	ユウ	amusement	友好度が上がる。(4.6)	2	normal (45.62)	normal (44.26)	「怒」→「怒」	「怒」→「」	0
217	ユウ	カナ	amusement	友好度が上がる。(2.29)	4	normal (47.89)	normal (55.62)	「怒」→「怒」	「」→「喜」	0
229	ユウ	カナ	way	友好度が上がる。(8.01)	6	normal (47.9)	normal (54.5)	「怒」→「怒」	「喜」→「喜」	0
296	ヒロ	カナ	lectureC	友好度が下がる。(−1.72)	8	normal (41.26)	normal (44.77)	「」→「」	「喜」→「喜」	0
297	カナ	ヒロ	lectureC	友好度が上がる。(1.41)	9	normal (43.04)	normal (41.26)	「喜」→「喜」	「」→「」	0
396	ユウ	ヒロ	amusement	友好度が上がる。(14.95)	10	normal (58.54)	normal (52.88)	「」→「」	「」→「楽」	0
397	ヒロ	ユウ	amusement	友好度が上がる。(12.61)	11	normal (53.76)	normal (50.09)	「楽」→「楽」	「」→「」	0
536	カナ	ヒロ	lectureB	友好度が上がる。(13.14)	17	normal (44.11)	normal (51.51)	「喜」→「喜」	「楽」→「楽」	0
538	ヒロ	カナ	lectureB	告白して成功する。	21	normal (51.51)	normal (44.11)	「楽」→「楽」	「喜」→「喜」	0

図 1: 生成されたストーリーベクトルの例

表 1: 実験条件

$x_{\max} \times y_{\max}$	120 × 120
エージェント数 $n$	3
エージェント名	ヒロ, カナ, ユウ
性別	男, 女, 男
最大ターン数 $t_{\max}$	6000
閾値 $h$	15

に関して、被験者には 1 つのストーリーベクトルからストーリーとなる文を 1 文生成するよう制限した。また、被験者の負担軽減のため、ストーリーベクトルの生成数が 10 以下の結果を、人間が理解しやすいように変換し利用した。実験は 2 回行った。図 1 に提案手法を用いて生成されたストーリーベクトルの例を示す。

**結果と考察** 実験後の聞き取り調査において、ストーリー生成の難易度を 1 (易しい) から 5 (難しい) の 5 段階で評価してもらった結果の平均はそれぞれ A が 3.5, B が 4.0 であった。このことから、提案手法を用いたストーリー生成は短編小説を書いたことがある人間にとっても困難であることが分かった。その原因については、ストーリーベクトルに各登場人物の性格や小説における役割、登場人物間の関係性について明示されていなかったためだと考えられる。これは、聞き取り調査において登場人物のキャラ設定や関係性を把握してからストーリーを生成するという意見を得ていることから確かである。

### 3.3 実験 2

実験 1 で生成してもらったストーリーを用いて実際に小説を書いてもらった。小説の生成に関して、必要最低限の要素で小説を書いてもらうためにストーリー 1 文に対して 140 字以下という制限を与えた。実験 1 と同様、実験は 2 回行った。

**結果と考察** 実験 1 と同様に小説生成の難易度を 5 段階で評価してもらった結果、それぞれの平均が、A が 1.5, B が 2.0 であった。このことから、自身

が考えたストーリーに沿って小説を生成することは、人間にとってそれほど困難ではないと考えられる。生成方法に関しては、聞き取り調査からストーリーの前後関係を補完するために文を追加したり、ストーリーに対して理由づけを付加することで文章全体の繋がりを強化していることが分かった。また、文の前後関係に合わせて使用する動詞や表現方法を適当に変化させていることも分かった。

## 4 まとめと今後の課題

本研究では既存の小説を利用しない小説の自動生成を最終目標として、既存の小説を利用しないストーリーの半自動生成手法を提案した。また、実際に提案手法を用いて人手で生成した小説を解析することで小説の自動生成に必要な要素を考察した。

今後の課題として、執筆経験がある人間のみだけでなく、執筆経験のない人間を含めた多数の被験者に実験し、小説の生成手法に対する更なる知見を得ることが挙げられる。

なお、本研究は一部、日本学術振興会科学研究補助金基盤研究 (C) (課題番号 26330282) の補助を得て行われたものである。

## 参考文献

- [1] 松原仁, 佐藤理史, 赤石美奈, 角薫, 迎山和司, 中島秀之, 瀬名秀明, 村井源, 大塚裕子: コンピュータに星新一のようなショートショートを創作させる試み, 第 27 回人工知能学会全国大会, 2D1-1 (2013)
- [2] 香山卓, 角薫: 星新一の物語を利用したユーモアのある物語生成システムの考察, 第 28 回人工知能学会全国大会, 1C2-OS-14a-4 (2014)
- [3] 今渕祥平, 小方孝: プロップ理論を包括するストーリー生成機構の開発の現状と課題, 第 28 回人工知能学会全国大会, 2F4-OS-01a-3 (2014)