

台詞に基づく個性を考慮した話者の分散表現に関する考察

佐藤 幸一 福田 清人 森 直樹 松本 啓之亮

大阪府立大学 工学研究科

1 はじめに

近年、小説や新聞、対話エージェントの発話文など、様々な文章を自動生成する試みが自然言語処理の分野でなされており、その際に、自動生成した文の登場人物に何らかの人物像を与え個性的な発話をさせることが注目されている。物語の登場人物は何らかの人物像を持っており、それは登場人物の台詞によって表現され我々に強く印象付けられると仮定し利用を考える。この場合、任意の文の特徴箇所を言い換えることで別の人物像を付与された台詞が自動生成できると期待される。

上記内容の実現のためには、台詞における人物像の特徴抽出や、言い換えの元文には別の人物像が表れていないか、言い換え後に求める人物像が付与されているかの判定が必要となる。単語や文の分散表現の獲得については積極的に研究されているが、話者そのものについての分散表現獲得については十分な研究がなされていない。そこで、本研究では台詞からその話者を推定し、話者の分散表現を生成する手法を提案する。

以下に本稿の構成を示す。2章で提案手法について説明し、3章で提案手法による実験について述べ、4章がまとめと今後の課題となる。

2 提案手法

本研究で提案する台詞における話者推定および話者の分散表現生成手法について述べる。提案手法では機械学習手法としてニューラルネットワークを用いる。

2.1 従来手法

本研究と目的に近い研究として、seq2seq と転移学習を用いて特定の人物像を付与した対話応答の生成を目的とした研究 [1] や、書き換えによって発話文に人物像を埋め込む際に必要な表層文の自由度について調査した研究 [2] がある。本研究の従来研究と異なる点は、まったく学習していない語彙を含む文に対しても

話者の推定している点と、人物像という定性的なものについて、分散表現を生成することで定量化した点である。

2.2 Word2Vec の学習による台詞の分散表現化

話者推定用のネットワーク学習前に分散表現への変換用のネットワークを Word2Vec [3] を用いて学習する。ここで獲得したネットワークの入力層から中間層を通すことで、台詞を任意次元の分散表現に変換する。これを話者推定用のネットワークの入力として用いることで、台詞内の使用単語を直接入力として用いる場合に比べて類義語についての学習が均一化されることや、似た内容について話す人物を人物像が近いとして学習されること、特定の単語を使っているかだけで話者を識別するように学習してしまう問題を回避することが期待できる。

以下に提案手法の学習アルゴリズムを示す。

1. 学習に用いる台詞集合を T 、 T の各台詞に対応した話者のラベル集合を L とする。
2. T を学習データとして、台詞の分散表現生成用ネットワーク C を Word2Vec で学習する。
3. T の各台詞 t について形態素解析で単語に分割、BoW を C に入力し中間層でのベクトル v_t を台詞の分散表現として取得する。
4. 3 で取得した v_t から T に対応した分散表現集合を V_t とする。
5. V_t と L の対応した各要素を入力、正解ラベルとして話者推定用ネットワーク P を学習する。

2.3 台詞による話者推定および話者の分散表現生成

話者推定および話者の分散表現生成には 3 層のニューラルネットワークを用いる。未知文の話者推定を

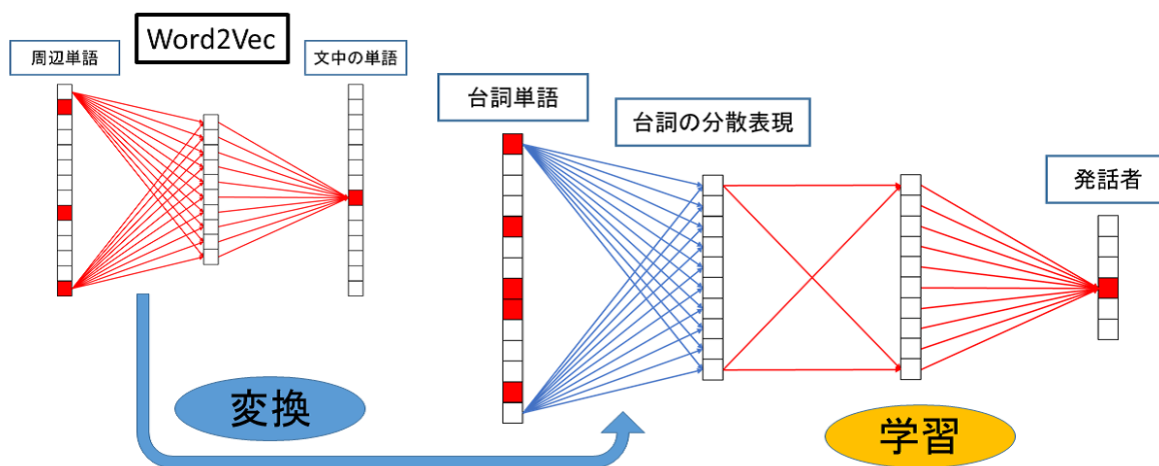


図 1: 提案手法の概要図

したい場合 2.2 節で学習済みのネットワーク C を用いて同様に入力文の分散表現を取得し、話者推定用ネットワーク P により未知文の話者として学習した人物のうち誰に適しているかを推定する。また学習によって得られた中間層から出力層への重みを Word2Vec における単語の分散表現のように話者の分散表現として取得する。

図 1 に提案手法で用いるネットワークの概要図を示す。また、比較手法として単純に台詞の BoW を直接入力とした場合の話者推定および話者の分散表現生成も実験した。

一般的に話者は、年齢、性格、性別、社会的地位など複数の属性で表現可能である。もしも、そのような属性が発話に影響を及ぼしているのであれば、話者としては異なるが、共通する属性を持つ話者同士の分散表現には何らかの相関があることが期待できる。

3 実験

3.1 実験 1

提案手法により、小説中の台詞の話者推定をした。学習とテストで別の作品の台詞を用いることで、未知文への識別精度を測った。本研究では小説 3 作品を用い、小説 1 から 3 名、小説 2 から 3 名、小説 3 から 2 名の計 8 名分の登場人物の台詞を抽出・使用した。以降、人物を順に人物 0~7 と呼ぶ。なお、人物 0, 2, 3 は女性、人物 1, 4, 5 は男性である。小説 1 の人物 0, 1 と小説 3 の人物 6, 7 はそれぞれ口調や性格等の人物像が近く、人物 2, 5 は他のどの人物とも

まったく異なる口調である。以下に実験に用いた台詞の条件を示す。

1. 本文中に登場する、“ \square ”に囲まれた部分である。
2. 発言者以外の人物の発言が台詞内で引用されている場合除外する。
3. 5 単語以上である。

学習では人物 0, 1, 2 の 3 名、人物 0, 1, 2, 5 の 4 名、人物 0, 1, 2, 3, 4, 5 の 6 名の台詞を、テストでは人物 6, 7 の 2 名分の台詞を用い、正解ラベルとして学習時にはそれぞれの台詞の話者を、テスト時は口調が近しいため、人物 6 にはラベル 0、人物 7 にはラベル 1 を用いた。

表 1 に実験条件を、図 2 に学習に用いた台詞数と語彙数を示す。各ネットワークのパラメータはグリッドサーチによって決定した。

3.2 実験 1 の結果

図 2 に比較手法と提案手法での識別精度を示す。

学習対象の人物が 6 人のとき、提案手法で比較手法を上回る識別精度が得られた。学習対象が 3 人、4 人のときは学習に用いた人物の口調がすべて異なるため、どの単語を使っているかのみで未知の人物がどの口調に近いかわかり、比較手法がより高い識別精度を示していた。しかし、対象人物が 6 人のときに新たに追加した人物 3, 4 の 2 名は、学習に用いる人物 0, 1 およびテストに用いる人物 6, 7 と共通した特徴を持っているため、比較手法では共有する単語に引っ張られて識別を誤ったと考えられる。このことか

表 1: 実験 1 の条件

学習人物数	3, 4, 6
テスト人物数	2
テスト台詞数	444
比較手法のパラメータ	
ネットワークのノード数	(学習語彙数, 700, 6)
epoch	300
提案手法のパラメータ	
ネットワークのノード数	(500, 300, 6)
epoch	300
Word2Vec のパラメータ	
分散表現の size	500
window size	3
min count	1
epoch	200

表 2: 学習時の人物数に対する台詞数と語彙数

学習人数	3人	4人	6人
学習単語数	992	1306	2153
学習語彙数	2165	2708	3700

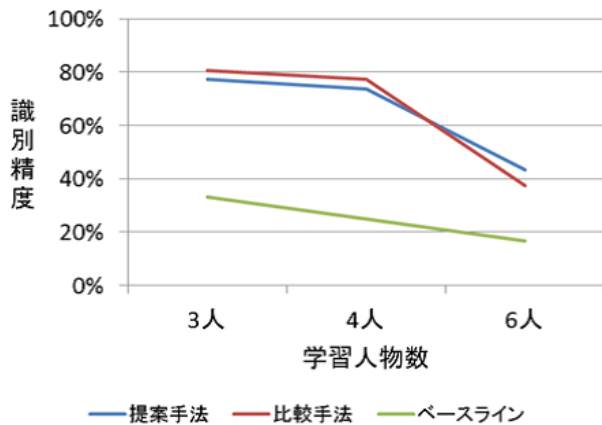


図 2: 識別精度

ら、提案手法は未知語を含む台詞の識別に有効だとわかった。

ここからは人物 0~5 の 6 名で学習した場合について述べる。表 3 にテスト時にそれぞれの正解ラベルだと識別されたデータ数を示す。なお、テストに用いた台詞の正解ラベルは人物 6 が 0, 人物 7 が 1 である。

提案手法と比較手法ともに、正解ラベル 0 の台詞はラベル 3 と、正解ラベル 1 の台詞はラベル 4 と誤識別されている場合が多かった。これは人物 0 と人物 3 の人物がそれぞれ女性かつ明朗な性格、人物 1 と人物 4 の人物がそれぞれ男性かつぶっきらぼうな性格のため、誤識別先としては妥当だといえる。また、比較

表 3: テストでの各正解ラベルへの識別数

手法	テスト人物	0	1	2	3	4	5
提案手法	6	82	13	1	144	10	23
	7	17	85	4	26	62	7
比較手法	6	117	1	4	94	13	14
	7	18	76	4	29	67	7

表 4: 実験 2 の条件

学習人物数	6
学習台詞数	2153
学習語彙数	3700
比較手法のパラメータ	
ネットワークのノード数	(3700, 50, 6)
提案手法のパラメータ	
ネットワークのノード数	(50, 6)
Word2Vec のパラメータ	
分散表現の size	50
window size	2
min count	1

表 5: SOM のパラメータ

入力データ	キャラクタ 6 人の分散表現
マップサイズ	(6, 6)
epoch	100000
近傍関数のコスト	0.001

手法では本来想定していた正解ラベル 0 よりもラベル 3 と識別されていた場合の方が多かった。これはテストに用いた人物 6 は口調・性格は人物 0 に近いが、“!” を多用し、人物 3 の台詞にも同様に“!” を含むものが多くみられるため、単語そのものを入力とする比較手法ではラベル 3 と識別してしまったものと考えられる。提案手法ではラベル 0 と識別された結果の方が多かったため、特定の単語を使っているかどうかで識別するような学習は抑えられていると言える。

3.3 実験 2

提案手法で生成した話者の分散表現を自己組織化マップ (Self Organized Mapping: SOM) [4] を用いて 2 次元に写像し、話者の人物像が表現できているか確認した。表 4 に実験条件を示す。epoch 数はすべて 200 である。

表 5 に SOM のパラメータを示す。また、SOM で写像するときの人物に振り分けた色は人物 0 が黄、1 が赤、2 が青、3 が緑、4 が紫、5 が水色である。

図 3, 4 に 比較手法及び提案手法で生成した話者

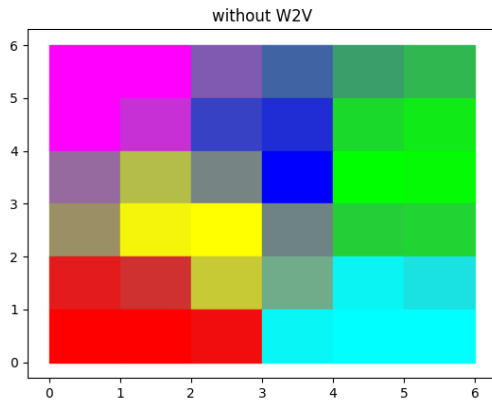


図 3: BoW を用いた話者分散表現の写像

の分散表現の写像結果を示す。台詞の BoW を入力とした既存手法は小説 1 の登場人物である人物 0 (黄), 1 (赤), 2 (青) が作品で固まって写像されているのに対し, Word2Vec による分散表現を用いた提案手法では作品ではなく共通した特徴を持つ人物 0 (黄) と 3 (緑), 人物 1 (赤) と 4 (紫) がそれぞれ近くに写像されている。このことから生成した分散表現は口調からキャラクター性を反映していると考えられる。

4 まとめと今後の課題

本稿では, 台詞の言い換えによる人物像の付与を最終目標として, 台詞における人物像の特徴抽出や人物像が付与されているかの判定のため今回は台詞による話者推定に主眼を置いた。特定の単語に依存した学習を防ぎ未知の語彙を含む台詞にも対応するため, Word2Vec による分散表現を用いた台詞の話者推定および話者の分散表現生成手法を提案し, 実験により未知台詞の識別が学習できていること, および話者の人物像を反映した分散表現が生成できていることを確認した。

今後の課題としては,

- 生成した話者の分散表現の解析。
- 台詞中の単語の時系列の考慮。
- 話者ごとに一律ではなく台詞ごとのラベル付けによる識別や, 台詞 1 つを 1 つの学習データとするのではなく, 話者で台詞をある程度まとめて 1 つの学習データとしての識別。
- 男性・女性や明るい・暗い性格のような話者に共通すると考えられる属性を正解ラベルとしての

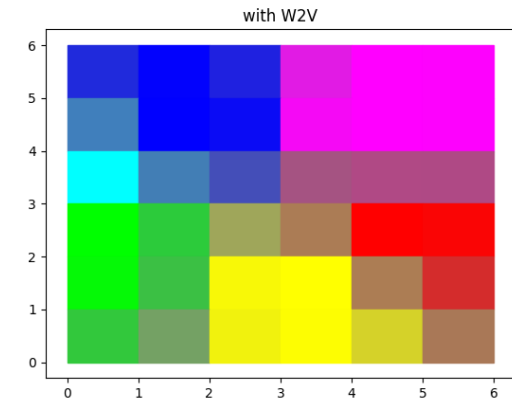


図 4: Word2Vec を用いた話者分散表現の写像

識別。

- 識別結果を用いての台詞の話者の人物像抽出および台詞の言い換えによる人物像の自動付与がある。

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] 赤間怜奈, 稲田和明, 小林颯介, 佐藤祥多, 乾健太郎. 転移学習を用いた対話応答のスタイル制御. 言語処理学会第 23 回年次大会発表論文集, 2017.
- [2] 刀山将大, 佐藤理史, 松崎拓也, 宮崎千明, 平野徹, 松尾義博. 文のどこにキャラクター性を埋め込む自由度があるか. 言語処理学会第 22 回年次大会発表論文集, 2016.
- [3] Quoc V. Le and Tomas Mikolov. Distributed representations of sentences and documents, 2014. <https://arxiv.org/abs/1405.4053>.
- [4] Willshaw and Von Der Malsburg. How patterned neural connections can be set up by self-organization, 1976.