観光案内におけるニューラル言語モデルを用いた説明文の生成

Language generation system for sightseeing guidance using neural language model

生田 和也* ¹	品川 政太朗 *1	吉野 幸一郎 *	¹ 鈴木 優 *1*2	中村 哲 *1*2
Kazuya Ikuta	Seitaro Shinagawa	Koichiro Yoshino	Yu Suzuki	Satoshi Nakamura
*1本白生神利受甘冻-	上学院士学 桂恕も	学 苮 応 彩 *2	デーク取動刑サイ	イエンフ創造レンタ、
示反元htth子1X11/NAIST Graduate so	人子阮八子 旧和祁 chool of information scie	十子和开子山个十 ence	ノース脳動型リイ NAIST Data	エノス別担セノタ a science center

In sightseeing information navigation systems, the information presented by natural language has a potential to improve usability. Several systems tried to embed the informing contents in a prepared template for generating sentences which are useful for tourists, which is called a slot filling based method. However, it is difficult for the systems to generate diverse expressions and unseen patterns. To solve this problem, we propose a neural network based sentence generation method instead of using a slot filling based method. In this research, we construct the contents as a one-hot vector representation and construct the neural network based language generator and the one-hot content vectors for generating natural and understandable sentences. We collected a tourist information corpus via crowdsourcing. Existing language generation systems used word classes. However, these systems often connect words unnaturally. In this research, we also proposed a re-ranking system based on a neural language model to solve the problem. In our experiments, we confirmed the naturalness and validity of the sightseeing guidance sentences generated by our proposed method.

1. はじめに

旅行者が観光プランを決定する上で必要以上の情報が溢れ, 逆に意思決定を困難とする問題を「情報過負荷状態」[Barta 09] と呼ぶ.この問題に対し、観光案内システムの開発では位置情 報や天候・混雑情報といったコンテキストを用いて提供する情 報を絞り込むのが一般的である [Abowd 99].この際,どのよ うな情報形式で提案を行うのかが重要となる. [Venkataiah 08] では、Web サービスなどで提供される様々なモーダルによる断 片的な情報を「Discrete Visualization」と定義し、それらを 単一モーダルとして集約した「Continuous Visualisation」と 比較することで、単一モーダルによる情報提示のほうが理解を 得やすいという結果が得られたと論じている.単一モーダルに よる情報提示のひとつとして,言語によるコンテンツの集約が 考えられる. 例えば観光者がルート案内システム [Tumas 09] により移動する時、臨時休業や運行ダイヤの変更によるルート 変更はしばしば起こり得る.この際,ルートのみによる情報提 供において、何故変更がなされるのかについての補足を行うの は難しい.これに対し、言語による情報提示では理由も含めて 提示を行う事ができ、ユーザの疑問を解消することができる. また情報提供を行うシチュエーションは多様に存在するのに対 し、モバイルインタフェースでは表示可能な画面の大きさやア クセシビリティが限られている。こうした状況に対し、言語に よる情報提示では、Push 通知や音声ガイダンスなどの情報提 供が可能である.本研究では、言語による情報集約を用いた観 光案内システムを想定し、伝達したいコンテンツが断片的に定 まっている際の文生成を行うことで、その有効性についての検 証を行う.

2. 文生成手法

文生成(Natural Language Genaration: NLG)は, 非言 語情報や中間表現を基いて人間が理解可能な文を生成すること であり,文書要約 [Conroy 01] や機械翻訳 [Cho 14], 詩の生成 [Zhang 14],株価動向の記述 [Murakami 17][Aoki 16] など幅 広いタスクにおいて用いられる.

2.1 ルールによる文生成

NLG において出力は文であるが入力としては様々なものを とりうる。これをルールによって生成するシステムの例とし て、地理情報システム (GIS)の構造化データを入力としてメッ セージを生成してルート案内を行うシステム [Dale 03] や、一 定時間ごとに観測した時系列データを入力とし、海上予報文を 生成する研究 [Sripada 03] などがある.これらでは文を生成 する際の言語ルールを人手により設計しているため高コストで ある.



図 1: テンプレートマッチングによる文生成

言 語 ル ー ル を 自 動 構 築 す る 研 究 と し て [Kondadai 13][山崎 16] では,NLG をテンプレート抽 出およびテンプレートマッチングの問題として扱うことで自 然な文を生成することに成功している (図 1). テンプレート とは固有名称をクラス表現タグとして置き換えた文の骨組み である.この手法では言語ルールの設計コストを少量に抑え て最適化を行うことができる反面,表現できる文の幅がテン プレートに依存するといった問題があり,クラス表現の組み 合わせによっては候補となるテンプレートを見つけ出すこと が難しい.

2.3 言語モデルによる文生成

文生成の異なるアプローチとして,言語モデルを用いるものがある.言語モデルとは文 $s = x_{1:n}$ において「単語 x_t の生起確率はそれまで出現した単語 $x_{1:t-1}$ に従って決まる」という仮定をおいて文をモデル化したものであり,予測される単語の生起確率に従ってサンプリングを行う事で生成が行わ

	× 1.			•	
PLACE	SHOP	TREE	EVENT	TREE-STATE	SALE
四条	松栄堂	桜	講演会	見頃	限定販売
河原町	舞扇堂	紅葉	祭り	散る直前	割引セール
BEAUTIFUL	GOOD-VALUE	CONGRSTION	BEGIN	STAY	END
True/False	True/False	True/False	True/False	True/False	True/False

表 1: コンテンツと埋め込み単語の対応

れる.モデル構築は再帰結合を有するニューラルネットワーク (RNN)が用いられる事が多く、これをニューラル言語モデルと呼ぶ. RNN において内部にメモリセルを持つものを Long-Short Term Memory (LSTM) と呼び、学習において 保持する情報、忘却する情報、出力する情報を最適化する事で 学習が行われる.これにより、時刻 t において生成される単語 が k である確率は単語予測分布 y を用いて下記で表される.

$$y_{t=k} \simeq P(x_t = k | x_{1:t-1}) \tag{1}$$

特に、コンテキストを用いて生成文を条件付けるモデルを Encoder-Decoder モデルと呼び、機械翻訳の研究 [Cho 14] な どで用いられる. [Cho 14] では翻訳元言語をコンテキストと して、翻訳先言語と対応付ける形で LSTM 言語モデルを構 築することで翻訳文を生成する. コンテキストを dとすると、 Encoder-Decoder モデルにおける予測分布 y は下記で表される.

$$y_{t=k} \simeq P(x_t = k | x_{1:t-1}, d)$$
 (2)

このアーキテクチャをそのまま本研究に適用しようとした場 合,コンテキストと説明文の対応関係が膨大となる.そのた め、コンテキストをある程度汎化して表現することができる枠 組みが必要となる。



⊠ 2: SC-LSTM

対話システムの研究において Wen らが提案した Semantically conditioned LSTM (SC-LSTM) [Wen 15] は, LSTM に新たな読み込みゲートとよばれる機構を加えて応答文の生成 を行っている (図 2).発話するクラス表現を「1」,発話しな いクラス表現を「0」としたバイナリベクトルをコンテキスト *d*として与え,読み込みゲートから得られるパラメータを減衰 係数としてコンテキストを収束させることで同時最適化を行い 言語モデルを構築する.これによりコンテキストの状態に応じ た柔軟な生成が可能となり,予測分布 y は下記になる.

$$y_{t=k} \simeq P(x_t = k | x_{1:t-1}, d_{1:t-1}) \tag{3}$$

本研究では観光案内システムのコンテンツは伝達内容を0/1の バイナリで表現したコンテンツベクトルとして生成すること ができると仮定し,その出力をコンテキストとして用いる事で テンプレートに依存しない観光案内文生成を行う.生成にあた り,Wenらのアーキテクチャをもとに実装を行う.

観光案内コーパスの構築

文生成を行うにあたり、まず観光案内を行う範囲を定め、対応したコンテンツの設計を行う.次に、設計したコンテンツにしたがってコーパスの収集を行う.

3.1 コンテンツの設計

Wen らが提案する生成システムのデザインと観光案内で必要とされる案内内容から、表1に示す12種類のコンテンツベクトルをデザインした。この際Wenらの提案システムでは、コンテンツベクトルは状態の時系列的な変化を考慮していないが、観光案内ではこうした状態にも対応を行う必要がある。そこで本研究では、観測状態の開始や終了、継続を表現可能な「BEGIN」「STAY」「END」といったコンテンツの定義を新たに加えた、今回のコーパス収集では、観光案内を行う地域を京都とした.

設計したコンテンツは、「Place =四条」といった埋め込み単 語として得られるもの、「Beautiful = True」といった抽象的な 表現として表現されるものがある.この表のコンテンツを組み 合わせてコンテンツパターンの設計を行う.そのためにまず、 あらかじめ文として表現可能なコンテンツパターン 69 通りを 設計し、コーパスの収集を行った.

3.2 コーパスの収集

学習に用いる文の収集にあたってはクラウドソーシング^{*1} を用いた.収集作業を行うにあたり,まずコンテンツパター ンからランダムに1つパターンを選択する.パターンに含ま れるコンテンツは対応したランダムな埋め込み単語あるいは True/False フラグに置き換えられ,作業を行うワーカに指示 される.ワーカには,その指示を満たす形で観光案内を行う説 明文を1文で記入してもらった(表 2).これにより構築され たコーパスを用いてモデルの構築を行い,観光案内文を生成 する.

	表 2:	ワーカによる記入例	
 0.2.			

コンテンツパターン	もみじ, 見頃,BEGIN=True
ワーカによる記入	もみじは見頃を迎え始めています
コンテンツパターン	松栄堂, タイムセール,GOODVALUE=True
ワーカによる記入記入	松栄堂のタイムセールはお得です

4. 観光案内文の生成

観光案内文を生成するために,SC-LSTM を用いて文の候 補を列挙する生成フェーズを行う.その後,文候補から適当な 文を選択するリランキングフェーズを経て最終的な観光案内文 を出力する.生成文の有効性を検証するにあたり,関連研究で 述べたテンプレートマッチングによる手法をベースラインとし た.以下で両手法の生成プロセスについて述べる.

*1 クラウドワークス:https://crowdworks.jp/

4.1 生成フェーズ

生成フェーズでは、コンテンツパターンをベクトル表現と したコンテンツベクトルを入力として与え、SC-LSTM ジェネ レータを用いて文候補の生成を行う.

テンプレートマッチングでは、入力となるコンテンツパター ンと一致するテンプレートを探索し、文候補とする.一致する テンプレートが見つからない場合、レーベンシュタイン距離を 用いて最も近いパターンを持つテンプレートを用いる.

4.2 リランキングフェーズ

リランキングを行う際、埋め込み単語によって言い回しが変 化する可能性がある. 例えば (講演会) というイベントを文に埋 め込む場合、「開演された」という言い回しは自然であるが、「祭 り」というイベントを埋め込む場合、「開演された」といった表 現は不自然になる. リランキングを行った後に埋め込み単語を 当てはめる場合、この言い回しの変化を考慮することができな い.よって本論文における観光案内文生成では、SC-LSTM が 生成した文のクラス表現 (Entity) に埋め込み単語 (Phrase) を当てはめた言語モデルによるリランキングを行う事で,埋め 込み単語を考慮した文生成を行った (Phrase-reranker). ま た、言い回しを踏まえたリランキングを行うことによる自然 性・妥当性の変化について比較するため、クラス表現を埋め込 み単語に置き換えずにリランキングを行う場合においても同 様に生成を行った (Entity-reranker). リランキングに当たっ ては, 生成フェーズで得られた SC-LSTM エントロピに加え て、LSTM エントロピ、生成文に含まれるべきコンテンツの 不足数・剰余数を総コンテンツ数で割った SERR を用いて最 適化を行った.

テンプレートマッチングでは, Phrase-reranker の LSTM 言 語モデルによるコストに従ってリランキングを行う.

5. 評価実験

5.1 実験設定

モデルの学習には、クラウドソーシングにより収集したコー パス 9645 文とそれに付随するコンテンツパターン 69 種を用 いる. 69 種のうち無作為に抽出した 10 種のコンテンツパター ンとそれに対応する観光案内文 1405 文を未知の案内文として 除外し、残った 8240 文をトレーニング:バリデーション:テ スト= 7:1.5:1.5 となるようにコーパスを分割し、トレーニン グコーパスを用いて学習を行った.

生成文の評価は、テンプレートを用いた言語生成の先行研究[山崎 16] に習い、BLEU スコア及び主観評価により行う. BLEU は n-gram 一致率により参照文とどれだけ近しいかを 測るものであり、人間が作成した参照文との機械的な比較を行 うことで、言語生成が適切に行われているかについてのいく つかの指標と相関を持つことが指摘されている。今回は、参照 文はワーカにより記入された案内文を用いる. 主観評価では、 「自然性」と「妥当性」について「1~5」の5段階評価を行う とともに、各生成手法における「-2~2」の5段階の嗜好テス トを行った.

実験にあたって、「テストコーパスにおける説明文を参照文 としたもの」と「予め除外した案内文を参照文としたもの」の 2種類について文生成を行い、上述の評価を行った.以降では これらをそれぞれ「既知のコンテンツパターンによる文生成」 「未知のコンテンツパターンによる文生成」と呼ぶ.これによ り、学習データに存在しないパターンに対する頑健性を考察 する.

5.2 実験結果-BLEU スコア

BLEU スコアにおける評価結果を表 3 に示す. 評価結果について,以下のような傾向が見られた.

表	3: E	BLEU	スコ	ア
---	------	------	----	---

既知のコンテンツパターンによる文生成		
手法	BLEU スコア	
Template matching	<u>47.70</u>	
Entity-reranker	46.55	
Phrase-reranker	46.72	
未知のコンテンツパタ	ハーンによる文生成	
Template matching	37.99	
Entity-reranker	41.09	
Phrase-reranker	41.07	

- (i) 未知のコンテンツパターンによる BLEU スコアは、既知 のコンテンツパターンに比べて全手法が低下する結果と なった. Template matching では BLEU スコア大きく 低下したのに対し、他の手法のスコア減少は小さかった.
- (ii) 既知のコンテンツパターンにおいては Template matching による手法が最も高スコアであるが,他の手法と比較し てスコアの差は少なかった.
- (iii) 未知のコンテンツパターンについては、Template matching の BLEU スコアが最も低くなった。Entity-reranker が最も高いスコアであるが、Phrase-reranker のスコアと ほぼ同程度であった。

(*i*)~(*iii*) より,未知のコンテンツパターンが与えられた場合,提案手法はベースラインに比べて柔軟な文生成が行われていると考えられる.

5.3 実験結果-主観評価

主観評価における結果を表4,表5に示す.

既知のコンテンツパタ	マーンによ	る文生成
手法	自然性	妥当性
Template matchng	<u>4.01</u>	4.20
Entity-reranker	4.00	4.22
Phrase-reranker	3.96	<u>4.24</u>
未知のコンテンツパタ	マーンによ	る文生成
Template matching	3.98	3.10
Entity-reranker	4.00	<u>4.10</u>
Phrase-reranker	3.75	3.94

表 4: 主観評価-自然性・妥当性

表4は、生成文を「自然性」及び「妥当性」について「1~ 5」で5段階評価されたものである.表5は、3つの生成手法 のうちそれぞれ2つについて、どちらが好ましいかを「-2~2」 で5段階で評価したものである.主観評価結果から、提案手 法による文生成の自然性・妥当性は3.75~4.24と高い結果と なり、概ね適切な文が生成できていると言える.

既知のコンテンツパターンによる文生成においては,各手 法の自然性・妥当性及び嗜好性全てにおいて BLEU スコアに よる評価と同一の結果が得られた.

未知のコンテンツパターンによる文生成では、自然性・妥当 性及び嗜好性において Phrase-reranker による手法が Entityreranker による手法に比べて劣る結果となった. Phrasereranker で用いられる LSTM 言語モデルは埋め込み単語が 含まれるため学習難易度が相対的に高く、リランキングが適切 に行われない場合があった事が要因だと考えられる.実際にい くつかの出力文では、同じ内容を複数回繰り返すといった不適 切な文がリランキングされている例が確認された.

MUNH V) -		, ,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,,				
Dí	Template	Entity	Phrase			
Pret	matchng	reranker	reranker			
Template						
matchng	-	0.026	0.041			
Entity			0.010			
reranker	-0.026	-	0.013			
Phrase	0.041	0.010				
reranker	-0.041	-0.013	-			
L hu o		1 - 1 - 1	Let D			
未知のこ	コンテンツパタ	マーンによる	文生成			
未知のこ	コンテンツパタ Template	マーンによる Entity	文生成 Phrase			
未知の Pref	コンテンツパタ Template matchng	マーンによる Entity reranker	文生成 Phrase reranker			
未知の Pref Template	コンテンツパタ Template matchng	マーンによる Entity reranker	文生成 Phrase reranker			
未知のこ Pref Template matchng	コンテンツパク Template matchng -	マーンによる Entity reranker -0.561	文生成 Phrase reranker -0.332			
未知のこ Pref Template matchng Entity	コンテンツパク Template matchng -	マーンによる Entity reranker -0.561	文生成 Phrase reranker -0.332			
未知のこ Pref Template matchng Entity reranker	コンテンツパタ Template matchng - 0.561	マーンによる Entity reranker -0.561 -	文生成 Phrase reranker -0.332 0.232			
未知のこ Pref Template matchng Entity reranker Phrase	コンテンツパタ Template matchng - 0.561	マーンによる Entity reranker -0.561 -	文生成 Phrase reranker -0.332 0.232			

表 5: 主観評価-嗜好

表 6: 未知のコンテンツパターンによる生成結果

パターン (A)	EVENT=祭り, CONGESTION, START
Template matchng	平等院では祭りが行われ、混雑しはじめています。
Entity reranker	祭りで混雑が始まっています。
Phrase reranker	祭りは混雑し始めています。
パターン (B)	PLACE=上賀茂神社, EVENT=特別展
Template matchng	上賀茂神社では特別展が行われていて混雑しています。
Entity reranker	上賀茂神社ではお特別展が行われています。
Phrase reranker	上賀茂神社では特別展が開催されています。

5.4 生成例

未知のコンテンツパターンによる生成例を表6に示す. Template matching による手法では,指示したコンテンツパ ターンを満たさない文生成が行われている。例えば(A)では, 指示していない「平等院」という埋め込み単語が現れている。 (B)では,指示していない「CONGESTION」(混雑)という 表現が含まれている。Entity-reranker, Phrase-reranker の両 手法では,多くのコンテンツパターンにおいて未知のパターン に対応した文生成が行われており,コーパスに存在しない観光 案内文の表現が可能となっている。また Phrase-reranker によ る手法では,いくつかの文において埋め込み単語による言い回 しの変化に対応した生成が行われている事を確認した。例え ば表6パターン(B)では、Entity-rerankerで「お祭り」とい う表現に対して「特別展」を誤って埋め込んだ「お特別展」と いった出力がされている。Phrase-reranker ではこのような生 成は見受けられなかった。

5.5 結言

本論文では、観光案内システムにおいて文による情報提示 を行うシステムを提案し、伝えたいコンテンツが得られている 時の文生成について検証を行った.従来手法では、未知のクラ ス表現の組み合わせを考慮した文生成が難しいという課題が あったが、本研究では SC-LSTM アーキテクチャをベースに、 LSTM 言語モデルを用いてリランキングを行うことで、未知 のコンテンツの組み合わせを考慮した文生成において提案手法 の有効性を確認した.

今後の課題として、以下のようなものが考えられる.1つ目 に、コンテンツベクトルの設計をより具体化する必要がある. 実際に観光案内システムを開発するにあたって、より広域のド メインにおいて観光案内を行うことが想定され、そのため観測 するデータ及びそこから推定できるコンテンツについてより具 体化を行う必要がある.2つ目に、コンテンツの前後関係や重 複を考慮した文生成について検討する必要がある.コンテンツ の種類によっては生成する順番により意味合いが変化する場合 があるため、位置関係や記述回数を考慮した文生成が可能とな る枠組みが重要であると考えている.

謝辞

本研究成果は、国立研究開発法人情報通信研究機構の委託 研究により得られたものです。

参考文献

- [Abowd 99] Abowd, G., Dey, A., Brown, P., Davies, N., Smith, M., and Steggles, P.: Towards a better understanding of context and contextawareness, in *Handheld and ubiquitous computing*, pp. 304–307Springer (1999)
- [Aoki 16] Aoki, K. and Kobayashi, I.: Linguistic summarization using a weighted N-gram language model based on the similarity of time-series data, in *Fuzzy Systems (FUZZ-IEEE), 2016 IEEE International Conference on*, pp. 595-6011EEE (2016)
- [Barta 09] Barta, R., Feilmayr, C., Pröll, B., Grün, C., and Werthner, H.: Covering the semantic space of tourism: an approach based on modularized ontologies, in *Proceedings of the 1st Workshop on Context, Information and Ontologies*, p. 1ACM (2009)
- [Cho 14] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., and Bengio, Y.: Learning phrase representations using RNN encoder-decoder for statistical machine translation, arXiv preprint arXiv:1406.1078 (2014)
- [Conroy 01] Conroy, J. M. and O'leary, D. P.: Text summarization via hidden markov models, in *Proceedings of the 24th annual international* ACM SIGIR conference on Research and development in information retrieval, pp. 406–407ACM (2001)
- [Dale 03] Dale, R., Geldof, S., and Prost, J.-P.: CORAL: Using Natural Language Generation for Navigational Assistance, in *Proceedings of the* 26th Australasian Computer Science Conference - Volume 16, ACSC '03, pp. 35-44, Darlinghurst, Australia, Australia (2003), Australian Computer Society, Inc.
- [Kondadadi 13] Kondadadi, R., Howald, B., and Schilder, F.: A Statistical NLG Framework for Aggregated Planning and Realization., in ACL (1), pp. 1406–1415 (2013)
- [Murakami 17] Murakami, S., Watanabe, A., Miyazawa, A., Goshima, K., Yanase, T., Takamura, H., and Miyao, Y.: Learning to Generate Market Comments from Stock Prices, in *Proceedings of the 55th Annual Meeting* of the Association for Computational Linguistics (Volume 1: Long Papers), Vol. 1, pp. 1374–1384 (2017)
- [Sripada 03] Sripada, S., Reiter, E., and Davy, I.: SumTime-Mousam: Configurable marine weather forecast generator, *Expert Update*, Vol. 6, No. 3, pp. 4–10 (2003)
- [Tumas 09] Tumas, G. and Ricci, F.: Personalized mobile city transport advisory system, Information and Communication Technologies in Tourism 2009, pp. 173–183 (2009)
- [Venkataiah 08] Venkataiah, S., Sharda, N., and Ponnada, M.: A Comparative Study of Continuous and Discrete Visualisation of Tourism Information, pp. 12-23, Springer Vienna, Vienna (2008)
- [Wen 15] Wen, T.-H., Gasic, M., Mrksic, N., Su, P.-H., Vandyke, D., and Young, S.: Semantically conditioned lstm-based natural language generation for spoken dialogue systems, arXiv preprint arXiv:1508.01745 (2015)
- [Zhang 14] Zhang, X. and Lapata, M.: Chinese Poetry Generation with Recurrent Neural Networks., in EMNLP, pp. 670–680 (2014)
- [山崎 16] 山崎健史,吉野幸一郎,前田浩邦,笹田鉄郎,橋本敦史,舩冨卓哉,山肩洋子,森信 介 他:フローグラフからの手順書の生成,情報処理学会論文誌, Vol. 57, No. 3, pp. 849-862 (2016)