

NAIST-IS-MT1651006

修士論文

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

生田 和也

2018 年 3 月 14 日

奈良先端科学技術大学院大学
情報科学研究科 情報科学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士（工学）授与の要件として提出した修士論文である。

生田 和也

審査委員：

中村 哲 教授	（主指導教員）
松本 裕治 教授	（副指導教官）
吉野 幸一郎 助教	（副指導教官）
鈴木 優 特任准教授	（副指導教官）

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*

生田 和也

内容梗概

近年、IoT と呼ばれる現実世界の情報をネットワークを通じて観測・収集する取り組みがなされており、この観測情報をもとにリコmendを行う観光案内システムなどへの応用が期待されている。こうした観光案内を行う上で、自然言語による情報提示はユーザビリティを高めると期待されている。観光案内のようなシステムで自然言語を用いる場合、伝達内容をあらかじめ用意した文のテンプレートに埋め込むことが行われてきたが、表現の多様性や未知のエンティティパターンに対応した文生成を行うことが出来ないといった問題があった。そこで本研究では、伝達内容をベクトル表現として用い、ニューラル言語生成と同時最適化することにより、テンプレートに依存しない観光案内文の生成を行う。モデル構築には、クラウドソーシングによって作成した観光案内コーパスを用いた。また、言語生成に際しては一般にクラス化が行われていたが、本研究では表現の埋め込みと言語モデルによるリランキングを行うことで、クラス化によって生じる不自然さの解消を試みた。ここから学習したモデルが、意図した伝達内容を考慮した観光案内文が生成可能であるかや、未知のエンティティパターンに対応可能であるかについて分析・検討を行う。

キーワード

言語モデル, 言語生成, 観光案内

*奈良先端科学技術大学院大学 情報科学研究科 情報科学専攻 修士論文, NAIST-IS-MT1651006, 2018 年 3 月 14 日.

Language generation system for sightseeing guidance using neural language model*

Kazuya Ikuta

Abstract

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 crowd-sourcing. 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.

Keywords:

Language model, Language generation, Tourist information

*Master's Thesis, Department of Information Science, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1651006, March 14, 2018.

目次

図目次	v
表目次	1
第 1 章 緒言	2
1.1 研究背景	2
1.2 研究目的	3
1.3 本論文の構成	3
第 2 章 観光案内システム	4
2.1 関連研究	4
2.1.1 サービス提供方式	4
2.1.2 観光案内システムが伝えるべき情報	6
2.2 言語による観光案内	7
2.3 まとめ	7
第 3 章 文生成システム	8
3.1 関連研究	8
3.2 テンプレートを用いた文生成	10
3.3 言語モデルによる文生成	12
3.3.1 N-gram 言語モデル	12
3.3.2 ニューラル言語モデル	13
3.3.3 ベクトル表現を用いた条件付き生成	17
3.4 まとめ	21
第 4 章 観光案内説明文の生成	23
4.1 コンテンツベクトルの設計	23
4.2 クラウドソーシングによるコーパスの収集	25
4.3 SC-LSTM を用いた観光案内文生成（提案手法）	26
4.4 テンプレートによる観光案内文生成（ベースライン）	28

4.5	まとめ	29
第 5 章	評価実験	30
5.1	実験設定	30
5.1.1	コーパスの前処理	30
5.1.2	評価指標	31
5.2	結果	32
5.2.1	BLEU スコアによる比較	32
5.3	主観評価	35
5.4	事例分析	36
5.4.1	既知のコンテンツベクトルによる文生成	37
5.4.2	未知のコンテンツベクトルによる文生成	38
5.5	まとめ	39
第 6 章	結言	40
6.1	本論文のまとめ	40
6.2	今後の課題	40
謝辞		41
参考文献		42
付録		47
A	実施アンケート及び得られた解答	47
B	設定したコンテンツパターン	48
発表リスト		49

図目次

2.1	観光案内システムの構成 [1]	5
3.1	パイプラインアーキテクチャ	8
3.2	テンプレートベースの文生成	11
3.3	リカレントニューラルネットワーク	13
3.4	LSTM 概要	15
3.5	LSTM アーキテクチャ	16
3.6	Conditioned-RNN	18
3.7	SC-LSTM	19
4.1	文生成プロセス（提案手法）	27
4.2	生成プロセス（ベースライン）	28

表目次

4.1	想定する観測データ	23
4.2	コンテンツとフレーズの対応	24
4.3	ワーカによる記入例	26
4.4	共通表現への置き換え	26
5.1	BLEU スコア	32
5.2	主観評価-自然性・妥当性	35
5.3	主観評価-嗜好	36
5.4	既知のコンテンツベクトルによる生成結果	37
5.5	未知のコンテンツベクトルによる生成結果	38

第 1 章 緒言

1.1 研究背景

情報基盤技術の発展により，実世界の様々な情報をリアルタイムにセンシングし，インターネット上で情報を公開する取り組みがなされている．たとえばスキー場のようなウィンタースポーツ施設であれば，雪のコンディションを利用者に確認してもらうために動画のような形で公開したり，食事店であれば季節の料理やクーポン情報などをスケジュールのような形で公開する事が考えられ，Web メディアにおける情報発信量は年々増加している．これらの情報は有益である反面，ユーザは提供サービスごとに情報をつなぎ合わせて意思決定を行う労力を必要とする．このような，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る「情報過負荷状態」[2] が課題となっており，これに対してユーザの嗜好にフォーカスしたりコメントシステムの開発が盛んとなっている．

観光案内システムはコンテンツのリコmend [3] やルート案内 [4]，旅行計画 [5] など案内方式が多岐に渡り，情報発信は位置情報・画像情報・データフレームやその組み合わせなど様々なモダリティに分かれる．この際，自然言語による情報提示は観光案内におけるユーザビリティの向上を図る上で有用となる．例えば位置情報を用いたルート案内システムにおいて予期せぬ渋滞が発生した場合，言語による状況提示を行うことでルート変更を行う理由を説明できる．また複数のモダリティによる情報提供は単一のモダリティに比べて理解が難しいという報告もなされており [6]，自然言語による内容の集約及びテキストや音声による情報発信は，モバイル端末という限られた表示領域における効率的な伝達手段となり得る．

現在の観光案内システムにおいて，言語による情報提示は「伝える内容を数パターンに絞る」か「テンプレートと呼ばれる文の骨組みに内容を当てはめて言語化する」ことによって行われる．しかしながらこれらの手法において生成可能な文の多様性は低く，またドメインの拡張を行う上で人手によるルール設計にかかるコストが高いことが課題となる．これに対してニューラルネットワークを用いた言語モデル [7] を構築をすることで，テンプレートに依存しない文生成を行う事が可能である．本研究では，クラウドソーシングによって作成した観光案内コーパスを用いて，観光案内システムが伝えたいコンテンツを表すベクトル表現とニューラル言語

モデルを同時最適化することで、テンプレートに依存しない観光案内文の生成を行い、生成結果についての分析を行う。

1.2 研究目的

本研究における主眼は、観光案内システムにおいて提供されるコンテンツの言語による情報提示であり、ユーザに伝達したい内容は既に定まっていると想定し、動的に変化する内容を表現可能な文生成を行う。

文生成を行うまでのプロセスとして、初めに言語モデルを構築する際に必要となるコーパスの収集を行う。この際、観光案内システムは推薦内容を決定するにあたり必要となる情報をフレーズあるいは Yes/No で定義可能であるようなバイナリフラグとして得られていると仮定し、その情報をもとにクラウドソーシングを用いて観光案内文の収集を行う。これにより得られたコーパスを用いて学習した言語モデルにより文生成を行い、実際に伝達内容を含んだ文生成が行われているかを、機械翻訳などで用いられる BLEU やクラウドソーシングによる主観評価を用いて評価し、生成された事例の分析を行う。

1.3 本論文の構成

本論文では、まず近年提案されている観光案内システムについて述べ、言語による情報提示の有用性について論じる（2 章）。次に文生成の際に用いられるアプローチについて説明を行い、テンプレートを用いた文生成及び、近年用いられるニューラルネットを用いた言語モデリングにおける生成手法について言及する（3 章）。次に、今回収集した観光案内コーパス及びベクトル表現によるコンテンツの設計について述べる（4 章）。その後、コーパスを用いて、テンプレートベースによる文生成手法をベースラインとして本手法との結果の比較を行い、本手法が観光案内文の生成として妥当であるか、用いた学習データに過度に依存しない文生成が行われているかについて評価する（5 章）、最後に、総括と今後の展望について述べる（6 章）。

第 2 章 観光案内システム

本章ではこれまでの観光案内システムにおける関連研究について述べる (2.1 節). 次に, これらの観光案内システムにおいて言語による情報提示の有用性について論じる (2.2 節)

2.1 関連研究

旅行者にとって必要以上の情報が存在することを「情報過負荷」[2] と呼び, 近年の観光案内システムでは, ユーザにパーソナライズされた情報を優先的に提供することでこの問題に取り組んでいる. この際, システムがユーザに提供するコンテンツは, サービスにより様々に異なる [1].

2.1.1 サービス提供方式

図 2.1 は観光案内システムにおける一般的な構成である [1]. 観光案内の流れとして, システムは外的コンテキスト (ソーシャルデータ・天候情報等) やユーザコンテキスト (プロフィール・嗜好・過去のレーティング情報等) を組み合わせて, 推論エンジンを用いて適切なコンテンツの選定を行う. コンテンツの候補はコンテンツリポジトリと呼ばれるデータベースに含まれ, 各サービスは用途毎に特化したコンテンツリポジトリを所持する. 提供されると決定されたコンテンツは, サービス形態によって様々な形に加工され, モバイル端末を経てユーザに提供される. 推薦されるコンテンツは目的地周辺のレストランや娯楽施設などの POI(Point Of Interest) を用いるのが一般的であり, 例えばユーザによって登録された嗜好に基づいて観光都市のアトラクション (美術館, 協会など) のリコメンドを行うシステム [3] がある. リコメンドされるコンテンツはスコアリングによりランク付けを行い, Web サイトなどと似通った階層リストベースのインタフェースとして表現されるもの, マップベースによるコンテンツの提示を行うもの [8][9], サウンドやビデオ形式 [3], テキストや画像 [10] を用いるもの等様々なものがある. また上記の派生として, 旅行者の食事・宿泊等にフォーカスしてレストランやホテルなどを提案す

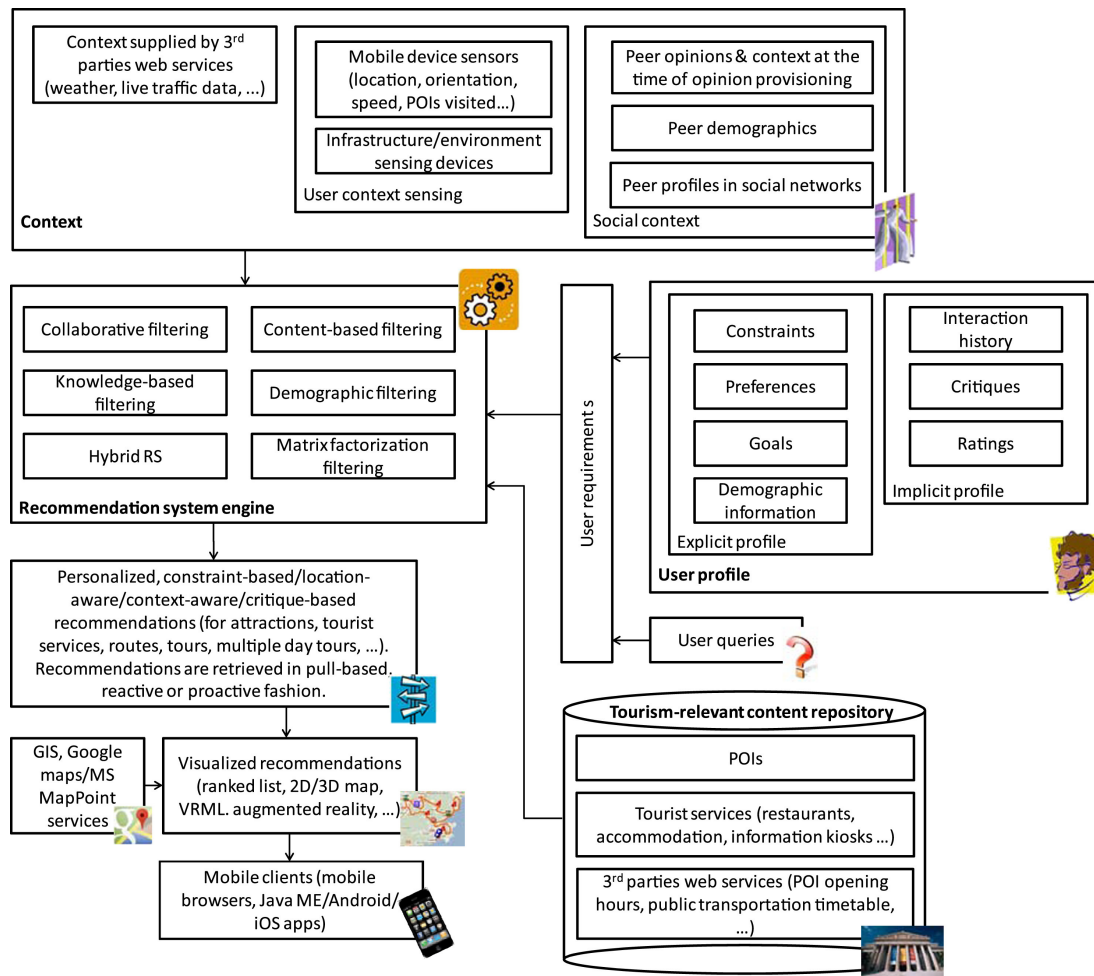


図 2.1 観光案内システムの構成 [1]

る観光案内システムも存在する [11][12][13].

単一的なスポットではなく観光におけるスケジュール・ルーティング自体を推薦する観光案内システムも存在する。例えばユーザに最も近いおすすめ POI への経路案内 [4], 交通乗り換え案内 [14] や経路案内に加えて旅行中の 1 日～数日間の観光スケジュールを推薦するシステム [15][5] などが提案されている。ソーシャルメディアを用いるものやユーザフィードバックを用いた観光推薦システム [16] では、地域で評判の良い観光スポットなどを見つけたり、トレンドを踏まえた推薦を行うことが可能となる。ソーシャルメディアにおいて提供される情報の価値は高く、他の旅行者の移動経路や口コミ等を用いた観光案内を行うシステムの研究が盛んに行

われている。

2.1.2 観光案内システムが伝えるべき情報

観光案内システムの意義は、サービス利用者に対して「状況に応じた有益となる情報を効率的に提供すること」である。効率的な情報提供を行うにあたっては、必要となる情報を推論するために外的コンテキストを利用するのが一般的であり [17]、外的コンテキストを用いる事で、観光中に予期せぬ事態が発生した際も、代替案の提案や状況説明を行う事が可能となる。例えば、観光客がすでに観光中である時、観光予定地が臨時休業であったり予定地周辺が混雑しているといった際に、空いている観光スポットの提案や迂回ルートの提案を行うことができ、観光プランのリスケジュールにおける意思決定の材料となる。外的コンテキストとしては、位置情報や天候・時間帯といった観測情報を用いられ [18]、モバイル端末ベースの観光案内システムでは、Web ベースの観光案内システムに比べて様々な外的コンテキストによる状況変化を取得した上での観光計画の提案が可能となる。

この際、どのような情報形式で提案を行うのかについてが重要となる。2.1.1 節で述べたような観光案内システムにおいて、コンテキストを用いたコンテンツの推論に焦点を当てた研究が盛んである反面、ユーザへのコンテンツ提供手段について踏み込んだ議論はあまり行われていない。モバイル端末におけるサービス提供インタフェースは Web ページを模した階層リスト表現やマップ情報によるものが多いが、必要な情報にアクセスするためのプロセスはサービスごとに異なっており、ユーザは提供サービス毎にシステムインタフェースの利用方法を把握する必要がある。Venkataiah ら [6] は、Web サービスにおいて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」としたコンテンツを作成し両者を比較することで、単一モーダルによる情報提示のほうが利用者の理解を得やすいという結果が得られたと論じている。例えば渋滞情報やイベント情報等が得られた時、これらの断片的情報を基に状況把握を行うためにはユーザ自身による推論ステップを介する必要がある、複数のコンテンツの統一的な情報提示は旅行者におけるユーザビリティの向上に寄与する可能性が高い。

2.2 言語による観光案内

単一モーダルによる情報提示の手段として、言語によるコンテンツの集約が考えられる。言語による情報提示において有用であると考えられる点として、以下の点が挙げられる。

一つ目に、言語による観光案内の補足を行うことができる点がある。例えば旅行者が [14] にあげられるようなルート案内システムにより移動するシチュエーションを考える。

ルート案内において、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。言語は我々人間が最も日常的に利用する情報伝達手段の 1 つであり、言語による情報提示は、ユーザの疑問を解消する手段の一つとして有用足り得る。また [15][5] 等の自動スケジューリングサービスにおいても、宿泊する施設の特徴や観光名所の言語によるナビゲーションは、純粋なコンテンツのみを用いた観光案内に比べて直接的な情報伝達の手段となる。

二つ目に、情報提供が可能なシチュエーションの多様性が挙げられる。モバイルインタフェースにおいて、視覚的な観光案内サービスにおける利用可能な表示領域は限られている。IoT をはじめとした観測情報の多様化に伴い、モバイル観光システムにおける外的コンテキストも増加していくことが予見され、このような観測情報を並べて提示するのはユーザビリティの観点から好ましくない。言語による情報提示では、これらの断片的な情報の集約を行える事に加え、Push 通知や音声ガイダンスなどのより多様な情報提供手段として利用することが可能となる。

2.3 まとめ

本節では、これまでに提案されている観光案内システムにおけるアプローチについて紹介し、統一的情報提示の有用性について論じた。言語による情報提示はモバイルインタフェースという限られた表示領域における多様な情報伝達手段となりえるが、観光案内において表現するコンテンツは天気や時間、イベント、交通情報、ユーザプロファイルなど多岐の組み合わせがリアルタイムに変化する。この際に、どのように言語化を行うかが重要となる。

第 3 章 文生成システム

本章では，まずこれまでに用いられている文生成システムにおける関連研究について述べ（3.1 節），テンプレートをベースとした文生成について述べる（3.2 節）．次に，近年注目されているニューラルネットを用いた文生成手法について述べ（3.3 節），先行研究における問題点及び本研究での取り組みについて述べる（3.4 節）．

3.1 関連研究

文生成（Natural Language Generation : NLG）は，非言語情報及び意味表現に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 [19] や機械翻訳 [20][21]，詩の生成 [22][23] など幅広いタスクにおいて用いられる．

文生成プロセスは大きく分けて「What to say : 何を言うか」と「How to say: どのように表現するか」からなっており，それぞれを分離して構築するパイプラインアーキテクチャ [24] を用いるのが一般的である．アーキテクチャの概要を図 3.1 に示す．

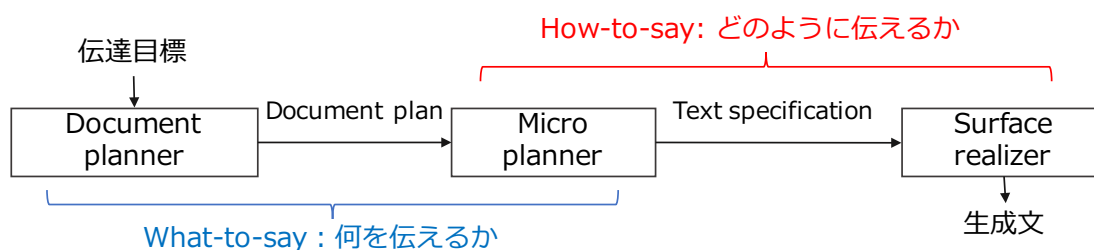


図 3.1 パイプラインアーキテクチャ

パイプラインアーキテクチャは伝達目標を持ち，ドキュメントプランナ，マイクロプランナ，リアライザと呼ばれる 3 つのモジュールを介して文生成を行う．「What to say」では，入力された非言語情報・意味表現等を基にドキュメントプランナを介して必要となる情報のフィルタリングを行い，マイクロプランナにより発話内容を表す中間表現にマッピングする．「How to say」では，得られた中間表現を基にリアライザを用いて文を生成する．本研究において，これは「観測データ，

ユーザコンテキスト等を用いた案内コンテンツの決定」と「決定されたコンテンツを用いた観光案内文の生成」と考えることができる。

NLG タスクにおいて出力は文であるが入力や中間表現の形式は定まっておらず、例えば言語によるルート案内を行うシステム [25] では、地理情報システム (GIS) の構造化データを入力としてマイクロプランナによって中間表現を作成し、リアライザの発話ルールを介して伝達メッセージの生成を行っている。海上予報文生成の研究 [26] では、一定時間ごとに観測した時系列データを入力とし、風向きや風速等のフレーズを中間表現として概要文を生成する。それ以外にも、将棋の盤面を入力に与えて得られた特徴単語を中間表現として状況説明文を生成するもの [27]、フローグラフと呼ばれる有効閉路グラフを中間表現として料理レシピ手順を表す文の生成を行うもの [28] などがある。

[25][26] などでは中間表現から文生成を行う際に使用される言語ルールを人手により設計している。言語ルールを手動で構築する場合は他のドメイン依存性が課題となり、ドメインの移行の際に再度ルール設計を行う必要がある。

人手による言語ルールの構築を行わずに文生成を行う手法も提案されており、修辭構造 [29] と呼ばれる文の構造がアノテーションされたコーパスを用いて文生成を行うもの [30] や、「属性-タグ」をペアとする中間表現をコーパスに付与して文生成を行うもの [31] などがある。多くの文生成においては、テンプレートと呼ばれる文の骨組みに中間表現を当てはめることで文生成が行われる [28][32][33]。

パイプラインアーキテクチャの特徴として、コンテンツの決定と文の生成をモジュール単位で分割している点が挙げられる。後の節で述べる Encoder-Decoder モデル [20][21][23] 等ではこのモジュールが分割されておらず、入力データと出力データの対応を直接マッピングすることで言語生成を行う。しかしながらこの方法による文生成を本タスクにおいて適用する場合、観測データとそれに対する説明文のペアを大量に用意する必要があり、その量は状況説明を行うドメインの広さにより更に増加する。また「What to say」において観光案内システムが推薦するコンテンツを決定するのに用いる入力データが変更される場合、そのデータを用いて新たなコーパスを収集し、最適化を行う必要がある。反面、コンテンツの推論と文生成を分離して構築することで、扱う観測データが変化したり観測するデータの種類を追加するような場合にも、最終的に伝達するコンテンツが変化しない場合におい

ては生成モジュールの再学習を行う必要が無い。また新たなコンテンツの追加が必要となる場合も、その場合に必要となるコーパスを収集して追加学習を行う事で生成モジュールの最適化が可能となる。コンテンツ推論に関しては既存の手法を用いて新たに文生成機能を付け加えるといったような機能毎のアップデートを行う事が可能であり、ドメインに依存しにくい枠組みであるといえる。

本研究では、上記に述べたパイプラインアーキテクチャによる文生成を想定する。そのため、伝えたいコンテンツは中間表現として既に既知であるとし、そのコンテンツをいかに文として表現するかに焦点を当てる。これはパイプラインアーキテクチャでいうところの「How to say」に当たり、中間表現は観光案内において伝えたいコンテンツを指す。これまでに述べたように文生成を行う手法はいくつか存在するが、中間表現を既知とする統計的文生成において、テンプレートを用いるもの(3.2節)と言語モデルを用いるもの(3.3節)がある。

3.2 テンプレートを用いた文生成

[28][32]らの手法では、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、従来手法に比べて自然な文を生成することに成功している。[28]ではフローグラフと呼ばれる有向グラフを中間表現として入力に与え、コーパスから生成したテンプレートを用いて言語モデルエントロピー等のコストに従ってマッチングし、フレーズを埋め込むことで文生成を行う。[32]では、構築したテンプレートに Culd と呼ばれる概念単位を付与し、各 Culd におけるテンプレート長やエンティティ数等を素性として用いてテンプレートマッチングを行い、フレーズを埋め込む事で文生成を行う。これらの手法における生成の流れを図 3.2 に示す。

テンプレートとは文を構成するための骨組みであり、学習コーパスから構築される。この際、場所や名前などのエンティティを共通のタグに置き換える事でテンプレートとなる。伝えたいコンテンツが中間表現として与えられる時、テンプレートリストのエンティティと中間表現のマッチング(テンプレートマッチング)を行い、同様のエンティティパターンを持つテンプレートが候補となる。その候補から最も適切なテンプレートを選択するために、テンプレート長やエンティティの位置・重

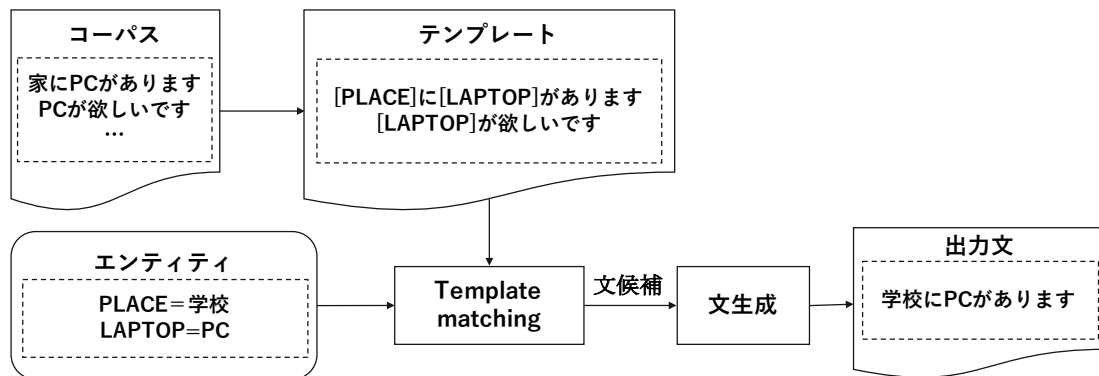


図 3.2 テンプレートベースの文生成

複率等を素性として選択が行われる。例えば [28] では、テンプレートの重複数やテンプレート内の内容語残存数、N-gram 言語モデルによるエントロピなどを選択コストとしている。テンプレートが決定された後、エンティティを中間表現に指定されたフレーズと置き換えて出力文となる。

この手法では、生成文は人間の記入した文をもとに出力されるため、言語ルールはコーパスから自動的に構築されたとみなすことができる。適用ドメインを拡張したい場合、対応するドメインのコーパスを収集することで、人手による再設計コストを少量に抑えて最適化することができる。反面、テンプレートを用いた文生成としていくつか課題が残る。1つは、テンプレートに存在しないエンティティパターンが中間表現として与えられた場合、候補となるテンプレートを見つけ出してマッチングすることが難しく、生成を例外なく行う場合には、考えられる生成パターン全てを網羅したコーパスの収集を行う必要がある。次に、中間表現が全てフレーズである必要がある。「地名」や「名前」といった固有的なものであれば固有辞書等を用いてエンティティとして共通化することが可能である。しかしながら「買う」「食べる」といった活用語尾の変化するフレーズをエンティティとして定義するのは手間を必要とし、エンティティとしたい表現が抽象的であるほど、テンプレートとして落とし込む難易度は上昇する。これらの点から、テンプレートを用いてフレーズに依存しない柔軟な文生成を行うのは難しく、テンプレートに依存しない文生成が求められる。

3.3 言語モデルによる文生成

コーパスベースの統計的文生成手法として、言語モデルを用いるものがある。テンプレートベースの文生成がフレーズを直接埋め込む手法であるのに対し、言語モデルベースの文生成ではコーパスを用いて言語モデルを構築し、言語モデルから得られる生起確率に従って単語をサンプリングし、つなぎ合わせることで文生成を行う。ここで述べた単語は「地名」や「名前」にとらわれない、フレーズ以外の全ての語彙である。

3.3.1 N-gram 言語モデル

言語モデルとは、文 $s = w_{1:n}$ において「単語 w_i の生起確率はそれまで出現した単語 $w_{1:i-1}$ 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、N-gram 言語モデルは特に「単語 w_i の生起確率が直前の N-1 単語にのみ依存する」という N 重マルコフ過程を用いるものである [7]。時刻 t において推定される単語 w_{t+1} は

$$P(w_{t+1}) = P(w_{t+1} | w_{t-(N-1):t}) \quad (3.3.1)$$

の条件付き確率で与えられ、この生起確率は

$$P(w_{t+1} | w_{t-(N-1):t}) = \frac{C(w_{t-(N-1):t+1})}{C(w_{t-(N-1):t})} \quad (3.3.2)$$

として求まる。ここで $C(w_{i:j})$ はコーパス内に単語列 $w_{i:j}$ が出現する頻度である。

$P(w_{t+1} | w_{t-(N-1):t})$ が高いほどコーパスによく出現しているパターンであるということであり、この生起確率に従って単語をつなぎ合わせることで文生成を行うことが可能となる。 $N = 2$ として構築したモデルをバイグラム (Bigram)、 $N = 3$ の時のモデルをトライグラム (Trigram) と呼ぶ。 N の数を大きく取るほどモデルのパラメータが増大し、ゼロ頻度問題 (Zero frequency problem) の発生する原因

となる．この対策として，単語の出現回数に一定数を加えた値を用いるなどのスムージング (Smoothing) を用いるのが一般的である．

3.3.2 ニューラル言語モデル

ニューラル言語モデルはニューラルネットを用いて文の生起確率をモデル化するアーキテクチャであり，**3.3.1** 節で述べた N-gram 言語モデルにおいて述べた N 重マルコフ過程で用いられる「 $N - 1$ 単語以上前の単語の影響を受けないという近似」を行わないものである．このニューラル言語モデルの構築を行う際には，一般的にリカレントニューラルネットワークがよく用いられる．

リカレントニューラルネットワーク (RNN) は，ニューラルネットワークでも特に再帰結合を有するものを指す．RNN を用いて言語をモデル化したものを RNN 言語モデル [34] と呼ぶ．アーキテクチャを図 **3.3** に示す．

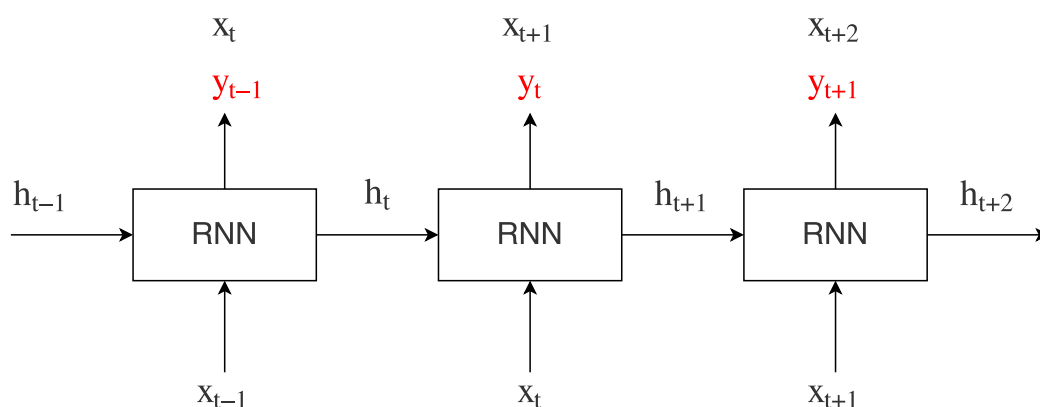


図 3.3 リカレントニューラルネットワーク

ここで h_t は時刻 t において得られる隠れ層と呼ばれる中間出力， x_t は単語 w_t を表すトークン， y_t は時刻 t におけるトークンの予測分布である．RNN 言語モデルの学習及び生成は，一般的に以下の手順からなる．

・学習データと教師データの準備

ニューラルネットワークは教師あり学習の一種であり，教師データを用意する必

要がある．単語トークン $T = x_1x_2...x_n$ が与えられた時，入力トークン列と出力トークン列

$$T_{in} = x_0x_1x_2...x_n \quad (3.3.3)$$

$$T_{out} = x_1x_2...x_nx_{n+1} \quad (3.3.4)$$

を用意する．ここで x_0 は開始を示すトークン， x_{n+1} は終了を示すトークンである．

・モデルの学習

各学習ステップにおいて T_{in}, T_{out} に対応するトークンを順番に入出力として与える．この時に予測される y_t の確率分布を出力トークン x_t に近づけるように最適化を行う事で学習が行われる．ここで，予測分布 y_t は，以下の式で導出される．

$$y_t = softmax(W_{hy}h_t) \quad (3.3.5)$$

$$h_t = tanh(W_{xh}x_{t-1} + W_{hh}h_{t-1}) \quad (3.3.6)$$

W_{ij} は重みパラメータ (Weight parameter) であり，これを調整することで最適化を行う． i, j はそれぞれ素性 (Feature) の変換前，変換後の次元サイズであり， $i \neq j$ のときこれは非線形変換となる．各ステップにおいて予測される中間出力 h_t は式 (3.3.6) より h_{t-1} と x_{t-1} の関数となっていることが分かる． h_{t-1} は更に前のステップ x_{t-2} と h_{t-2} の関数となっており，これによりイニシャルステップまでの情報が遡って伝搬され，理論的には過去全ての履歴に従ってモデル化が行われる．このため，分布 y_t において予測されるトークン $x_t = k$ である確率 $\hat{y}_{t=k}^{train}$ は以下で表す事ができる．

$$\hat{y}_{t=k}^{train} \simeq P(x_t = k | x_{1:t-1}) \quad (3.3.7)$$

・言語モデルによる文生成

文生成において，入力となるトークンは前ステップによって予測されたトークンとなる．よって生成ステップ t において予測されるトークン $\hat{y}_{t=k}^{test}$ の生起確率は，予測トークン x による条件付けが行われ，以下の式となる．

$$\hat{y}_{t=k}^{test} \simeq P(x_t = k | \hat{x}_{1:t-1}) \quad (3.3.8)$$

文生成においては，各生成ステップ毎に得られる予測分布 y_t に従って単語をサンプリングし，つなぎ合わせることで表現される．

言語モデルベースの文生成においてテンプレートのような文の雛形は存在しておらず，生成ステップ毎の単語の選び方によって様々な文の表現を行うことが可能である．この RNN 言語モデルの生起確率に従った文生成において公となっている問題として勾配消失問題 [35] が指摘されている．これは，生起確率の算出において参照する過去の履歴が深くなるにつれて情報が希薄になり，長期系列の履歴を踏まえた学習を行う事が難しいという問題である．

Long-Short Term Memory[36] と呼ばれるアーキテクチャでは，伝搬する情報を隠れ層に加えメモリセルと呼ばれる変数に保存して渡す事で，勾配消失問題を緩和することができる（図 3.4）．

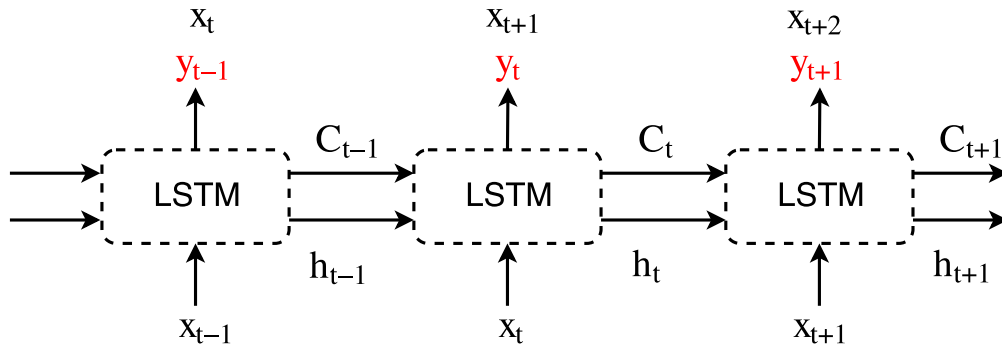


図 3.4 LSTM 概要

前述した RNN と異なる点として，伝搬するパラメータに C が加わっている点が挙げられる．これをメモリセルと呼び，内部状態の更新は積和のようなシンプルな演算のみを用いて更新される．図 3.5 は LSTM の内部アーキテクチャである．

図 3.3 で述べた RNN アーキテクチャに加えて，LSTM では入力ゲート i (Input gate)，忘却ゲート f (Forget gate)，出力ゲート o (Output gate) が新たに定義され利用される．ステップ t における各ゲートの状態は，以下の式で導出される．

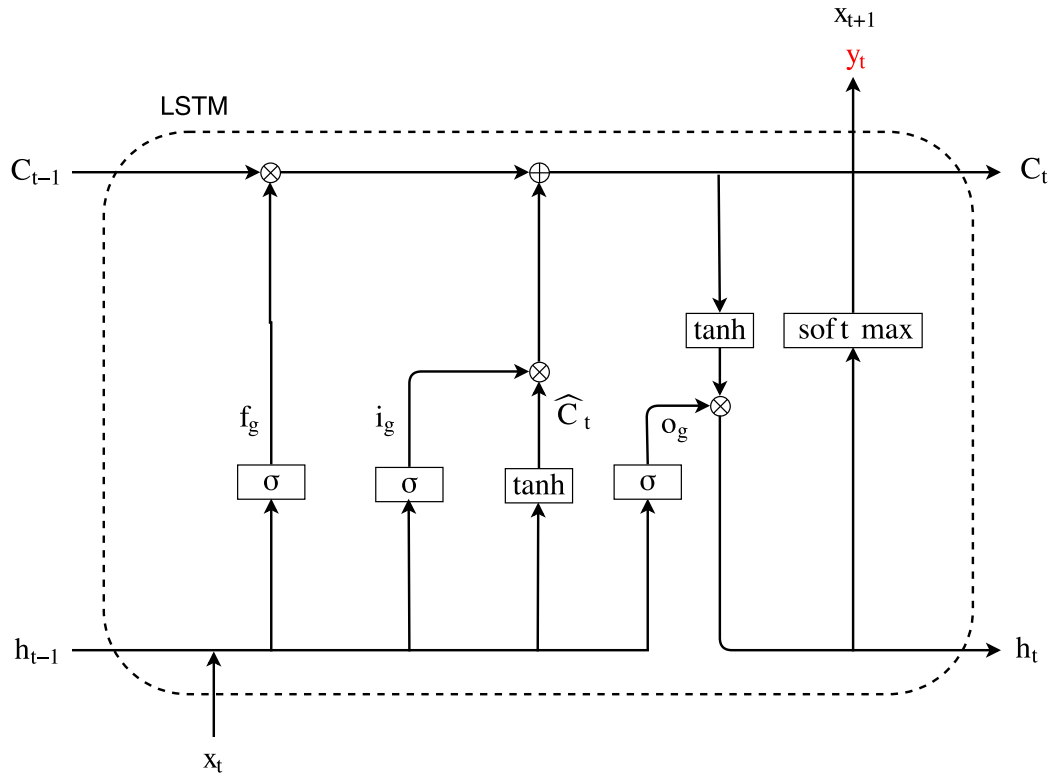


図 3.5 LSTM アーキテクチャ

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1}) \quad (3.3.9)$$

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1}) \quad (3.3.10)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1}) \quad (3.3.11)$$

上記ゲート出力を介して，メモリセルの状態 C_t は下記のように更新される．

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t \quad (3.3.12)$$

前ステップのメモリセルの内部状態は f_t を介して忘却され，今ステップにおけるメモリセルの状態は i_t を介して制御される．

最終的に得られる予測分布 y_t は O_t によって出力状態を更新された隠れ層 h_t を

パラメータとした以下の式で表される．

$$y_t = \text{softmax}(W_{ho}h_t) \quad (3.3.13)$$

$$h_t = O_t \odot \tanh(C_t) \quad (3.3.14)$$

LSTM は，ゲート機能により隠れ層 h_t 及びメモリセルの更新を行うアーキテクチャとなっており (3.3.12) により長期系列に対しての勾配消失に頑健な学習を行う事ができる．分布 y_t において予測されるトークン $x_t = k$ である確率 $\hat{y}_{t=k}^{train}, \hat{y}_{t=k}^{test}$ は式 (3.3.7), (3.3.8) と同様である．

3.3.3 ベクトル表現を用いた条件付き生成

Conditioned-RNN

3.3.2 節では，ニューラルネットを用いた文生成において，テンプレートを用いずに文を生成する統計的手法について述べた．ニューラル言語モデルでは，コーパスを基に言語ルールを自動学習することで生起確率に従った文生成を行うことができる．しかしながら，これまでに述べたアーキテクチャでは，単にコーパスをもとにそれらしい文を生成するに留まっており，伝えたいコンテンツに従った文生成を行うには，コンテキストとよばれる補足情報を付け加えて生成に条件付けを行う必要がある．これを条件付き生成と呼び，図 3.6 のようなアーキテクチャで表される．このアーキテクチャはメモリセル C を含まない RNN であり，メモリセルを有する LSTM などにおいては，隠れ層の状態 h に加えてメモリセルの状態 C が横方向に伝搬される．

条件付き生成では，入力にトークン x に加えて d が挿点される．この d をコンテキストと呼び，入力トークン x と連結して入力が行われる．予測分布は (3.3.5)(3.3.6) から以下に変化し，コンテキストに条件付けられる．

$$y_t = \text{softmax}(W_{hy}h_t) \quad (3.3.15)$$

$$h_t = \tanh(W_h[x_{t-1}; d] + W_{hh}h_{t-1}) \quad (3.3.16)$$

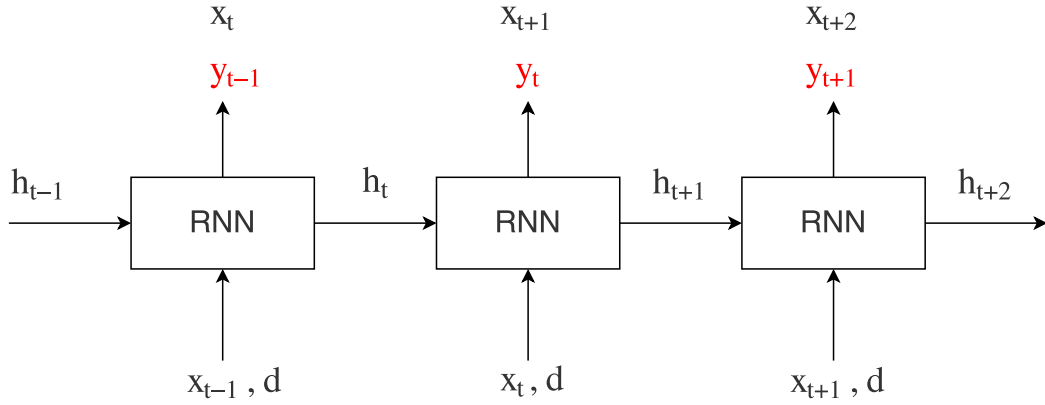


図 3.6 Conditioned-RNN

これにより, (3.3.7)(3.3.8) で示したステップ t における予測 $\hat{y}_{t=k}^{train}$ 及び $\hat{y}_{t=k}^{test}$ は, 以下のようにコンテキスト d により更に条件つけられる.

$$\hat{y}_{t=k}^{train} \simeq P(x_t = k | x_{1:t-1}, d) \quad (3.3.17)$$

$$\hat{y}_{t=k}^{test} \simeq P(x_t = k | \hat{x}_{1:t-1}, d) \quad (3.3.18)$$

コンテキストは柔軟に設定することが可能である. 例えば映画レビューサイトにおけるレビューと各レビュー評価 (星 1 ~ 5) をコーパスとして扱うことを考えた時, レビュースコア (1 ~ 5) をコンテキストとすることでレビュースコアに従った文生成を行うことができる.

Encoder-decoder

条件付き生成における応用例として Encoder-decoder がある. このモデルはシーケンスを異なるシーケンスに変換するものであり, 主に機械翻訳タスクなどに用いられる [20][21]. Encoder-decoder モデルでは, 条件付き生成におけるコンテキスト d は RNN エンコーダを介してエンコーディングされる. 例えば入力シーケンス T^{source} から異なるシーケンス T^{target} への変換を行う際, $T^{source} = x_{1:n}^{source}$ とすれば, コンテキストは以下で表される.

$$d = RNN^{encoder}(x_{1:n}^{source}) \quad (3.3.19)$$

この時、 T^{target} は RNN エンコーダによってエンコーディングされた d により条件付けられて学習され、変換が可能となる．このアーキテクチャは入力シーケンスと出力シーケンスの対応関係を直接学習する手法であり、パイプラインアーキテクチャとは異なる End-to-End な枠組みとなる．**3.1** 節で述べたように、観測データとそれに対する説明文の収集の難しさ、観測データの追加・変更による再学習などのコストを踏まえると、観光案内というタスクにおいては好ましくない．また観光案内においてコンテンツとなる情報は混雑情報や気温、天候など多岐に渡り、対応関係の学習が難しいことが想定される．

Semantically conditioned LSTM

Wen らが対話システムの応答文生成の研究において提案した Semantically conditioned LSTM (SC-LSTM) [37] アーキテクチャでは、コンテキストを動的に変化させる事でより柔軟な文生成を可能としている．アーキテクチャを図 **3.7** に示す．

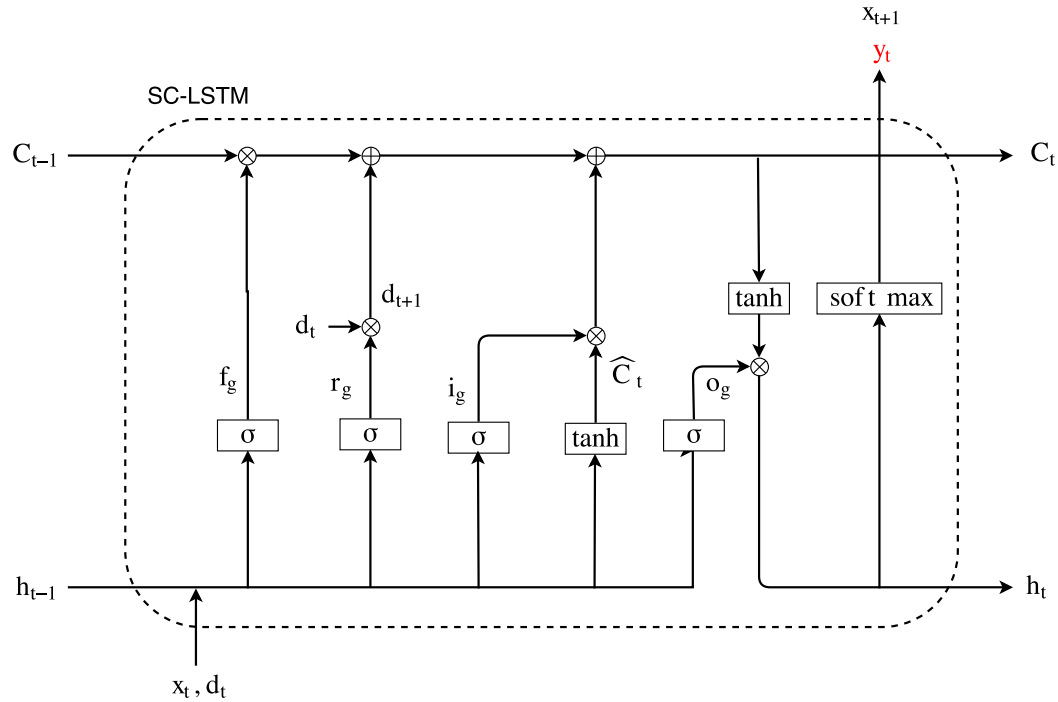


図 3.7 SC-LSTM

SC-LSTM が LSTM と異なっている部分として、新たな r というゲート機構が加えられている点がある． r は読み込みゲート (Reading gate) と呼ばれ、このゲートを介してコンテキスト d は動的に変化する． d の更新は以下の式で行われる．

$$r_t = \sigma(W_{xr}x_t + W_{hr}h_{t-1}) \quad (3.3.20)$$

$$d_{t+1} = r_t \odot d_t \quad (3.3.21)$$

(3.3.20)(3.3.21) より、読み込みゲート r はコンテキストにおける減衰係数 (Decay factor) とみなすことができる．このアーキテクチャにおいて、コンテキストは「1」か「0」で表現されるバイナリベクトルで定義され、このベクトルが最終的には全てが「0」となるように r の減衰係数を最適化することで、コンテキストはコーパスによって更に条件付けられる．

学習において、読み込みゲートの追加によりメモリセルの更新は式 (3.3.12) から式 (3.3.22) に変化し、バイナリベクトルの更新及び更新内容はメモリセルを介して次のステップに伝搬される．

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t + \tanh(W_{dc}d_{t+1}) \quad (3.3.22)$$

ステップ t における予測分布は以下の式で表される．

$$\hat{y}_{t=k}^{train} \simeq P(x_t = k | x_{1:t-1}, d_{1:t-1}) \quad (3.3.23)$$

$$\hat{y}_{t=k}^{test} \simeq P(x_t = k | \hat{x}_{1:t-1}, \hat{d}_{1:t-1}) \quad (3.3.24)$$

(3.3.23)(3.3.24) から、入力トークンとバイナリベクトルの状態がステップ毎に変化し、その両方によって条件付けられている事が分かる．

Wen らは、対話システムにおいて対話管理部が決定したシステム応答の中間表現をバイナリベクトル d としている．中間表現はバイナリベクトルと対応付けられており、バイナリベクトルの各インデックスは応答文において含まれているエンティティとして定義される．このエンティティはフレーズあるいは文に含まれる抽象的な表現である．フレーズは「Name」や「Area」といったエンティティ毎に共通表現に置き換えられて学習が行われる．これにより学習されたモデルを基に文生成を

行い、生成された文に対して再度言語モデルを用いてエントロピを算出し、生成文に含まれるフレーズの一致数とエントロピを用いてリランキングを行い、出力文とする事で、自然かつ妥当な文生成を可能としている。

3.4 まとめ

本章では、文生成において用いられる主な手法について述べた。文生成の枠組みとしては、伝達するコンテンツを決定する「What to say」と得られたコンテンツを文として表現する「How to say」を分離して構築するパイプラインアーキテクチャに従うものが一般的であり、近年では、機械翻訳などのタスクで用いられる入力データと出力データの対応関係を直接学習する End-to-End な手法も存在する。本研究では、観測データと説明文をペアとしたコーパス収集の難易度、観測情報を変更した時の文生成アーキテクチャの再利用性などの観点から、パイプラインアーキテクチャをベースとし、伝達したいコンテンツが既知である時の文生成について検討する。コンテンツが既知である時の文生成においては、テンプレートを用いた手法がある。この手法では、文の骨組みとなるテンプレートにフレーズを埋め込むことによって文生成を行うが、与えるフレーズの組み合わせによっては生成を行うことができないといった課題や、フレーズによらない抽象的なコンテンツを表現できないという特徴がある。

本研究では、Wen らが対話システムの応答文生成において提案したベクトル表現を「観光案内システムが提供したいコンテンツ」を表すコンテンツベクトルとし、言語モデルと同時最適化を行うアーキテクチャをベースとする事で、観光案内においてテンプレートに依存しない文生成を行う。ここで、観光案内というドメインにこのアーキテクチャを使用するにあたって、いくつか考慮すべき点がある。

まず、観光案内システムにおいて提供される情報は、混雑具合やイベントスケジュール等によって動的に変化する事が予見される。そのため、Wen らにより用いられるアーキテクチャを観光案内システムにおいて用いる場合、先行研究では考慮されていない「リアルタイムな説明文を生成可能となるコンテンツ設計」を行う必要がある。これに対して本研究では、観測した状態が「始まったのか」や「続いているか」などの時々状態をふまえたコンテンツ設計を行うことでこの問題に取り

組む。

またこれに付随して，伝達するコンテンツによって言い回しが変わる可能性が考えられる．例えば観光案内システムがイベント情報の通知を行いたい時，そのイベントが「講演会」であれば「開演した」という表現を使って情報提供を行うのは自然である．しかしながら，イベントが「お祭り」である場合，言い回しは「開演した」というより「始まった」といった表現であるほうが好ましい．Wen らの用いる文生成プロセスにおいて，リランキングに用いる言語モデルはフレーズをエンティティに置き換えた文を基に学習が行われており，埋め込むフレーズによって変化する接続等に対応することが難しい．本研究では，対話システムにおける応答文生成の枠組みが観光案内システムに適用可能であるかについて検証するとともに，エンティティを含まないコーパスを基に設計した異なる言語モデルを用いてリランキングを行う手法を提案し，それによりフレーズを考慮した文生成が可能であるかについて検討する．

第 4 章 観光案内説明文の生成

3 章で述べたアーキテクチャを用いて観光案内文の生成を行う上で、モデルの学習に用いるために観光案内を行っている説明文を収集する必要がある。本研究では、観光案内文の収集を行うためにクラウドソーシングを用いた。

本章では、まずクラウドソーシングを用いて収集するために設計したコンテンツベクトルについて述べる (4.1 節)。次にクラウドソーシングにおける観光案内文の収集方法について述べる (4.2 節)。次に、本研究における説明文生成の提案手法について述べ (4.3 節)、比較を行うベースラインについて述べる (4.4 節)。

4.1 コンテンツベクトルの設計

本研究では伝達したいコンテンツを「1」、伝達する必要の無いコンテンツを「0」とするバイナリベクトルを条件付き生成におけるコンテキストとし、これをコンテンツベクトルと呼ぶ。

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、本研究室及び関連研究室的観光案内システム関係者にアンケート (付録 A) を実施した。具体的には「実際にどのようなデータを観測可能か」「どのようなコンテンツが得られ、それによりどのような観光案内を行う事ができるか」といった質問を提示し、主に外的コンテキストの収集を担当する研究室の学生らに回答を依頼した。その結果、主に表 4.1 に述べる情報が収集されるという回答が得られ、それらを参考にできるだけ多くの意見を反映する形でのコンテンツ設計を行った。表 4.2 が設計したコンテンツの詳細である。

表 4.1 想定する観測データ

観測データ	得られる情報
動画像情報	桜や紅葉などの見所、見所の状態
ソーシャルデータ	トレンド情報
センシングデータ	混雑情報
Web クローリングデータ	店舗情報、イベントスケジュール

表 4.2 コンテンツとフレーズの対応

Place	Shop	Tree	Event	Treestate	Sale
四条	一保堂茶舗	桜	講演会	見頃	限定販売
金閣寺	嵐山茶屋	紅葉	祭り	散る直前	タイムセール
四条	柳桜園茶舗	もみじ	ミュージアム	七分咲き	割引セール
平等院	京極井和井		特別展		
平安神宮周辺	松栄堂		お花見		
河原町	舞扇堂				
上賀茂神社	美山荘				
二条城	竹屋町三多				
天満宮周辺	瓢亭本店				
	柳桜園茶舗				

Beautiful	Goodvalue	Congestion	Begin	Stay	End
True	True	True	True	True	True
False	False	False	False	False	False

設計したコンテンツは 12 種類あり，上段のコンテンツ「Place」「Shop」「Tree」「Event」「Treestate」「Sale」はフレーズ（例：Place = 四条）として得られ，下段のコンテンツ「Beautiful」「Goodvalue」「Congestion」「Begin」「Stay」「End」は True/False（例：Beautiful = True）などの多様なフレーズとして表現しうる状態で表される．

Wen らの提案した応答文生成システム [37] では，米国 SF サンフランシスコのレストランとホテルの対話を行うシステムにおいて用いられるエンティティ [38] に基いて設計されたコンテンツを使用している．これに対し本研究では，リアルタイムに変化する状況を説明可能とするようなコンテンツ設計を行うため，「Begin」「Stay」「End」といった「始まったのか」「続いているのか」「終わったのか」を示すコンテンツを設定している．これにより，より動的な状態変化を考慮した文生成を行う事が可能となる．この表をコンテンツテーブルとしてコンテンツパターンの設計を行う．

このコンテンツパターンは慎重に決める必要がある。例えば「Place = 四条, Treestate = 七分咲き」というようなコンテンツパターンから作られる文は「四条は七分咲きです」というような実際の観光案内では生じ得ない表現としかなりえない。これを避けるために、あらかじめ文として表現可能なコンテンツパターンを用意する。今回のコーパス収集では、意味のある文が記入可能なコンテンツパターン 69 通りを設計（付録 B）し、コーパスの収集を行った。このコンテンツパターンを用いてコンテンツベクトルを作成する。

コンテンツベクトルは、コンテンツパターンにおいて伝達したいコンテンツに対応するインデックスを「1」とし、伝達する必要の無いコンテンツを「0」としたバイナリベクトルである。コンテンツベクトルにおいて、「1」であるコンテンツは表 4.2 に紐付けられる。これにより、フレーズはエンティティとして定義される。True/False で表現されるコンテンツは、その表現を述べたい場合に「1」とすることで、抽象的な表現をコンテンツとして定義する。

4.2 クラウドソーシングによるコーパスの収集

4.1 節により設計したコンテンツパターンを用いてクラウドソーシング*でコーパスの収集を行う。ワーカには、コンテンツパターンからランダムに選んだパターンを用いて案内するコンテンツの指示を行う。指示は選ばれたパターンに従って表 4.2 からランダムにフレーズが選択される。True/False で表現されるコンテンツについては、含める必要がある場合のみに、「Beautiful」といったようなコンテンツの名前を提示し、それらにより観光案内を行う説明文を 1 文で記入してもらう。品質を確保するために、記述の前にいくつか練習問題を解いてもらい、それをクリアした上で記入された文の承認/非承認作業を行う。

指示するコンテンツパターン及びワーカにより記入された文例を表 4.3 に示す。表 4.3 において、上段が承認したアノテーション、下段が非承認としたアノテーションである。下段の例では「指示を行っていないコンテンツが含まれている」「1 文で記入を行うという指示が満たされていない」といった理由から非承認とした。それ以外にも、明らかな非文であるものや、指示を理解していないような記入に関

*クラウドワークス : <https://crowdworks.jp/>

して承認/非承認作業を行い，登録されているフレーズをエンティティタグに置き換える事でコーパスを構築する (表 4.4).

表 4.3 ワーカによる記入例

コンテンツパターン	ワーカの記入（承認）
もみじ 見頃 BEGIN	もみじは見頃を迎え始めています
松栄堂 タイムセール GOODVALUE	松栄堂のタイムセールはお得です
河原町 ミュージアム END	河原町のミュージアムは終了間近です
コンテンツパターン	ワーカの記入（非承認）
二条城 もみじ 見頃 限定販売 GOODVALUE	二条城のもみじは 11 月上旬が見頃 100 個限定販売。お得です

表 4.4 共通表現への置き換え

ワーカによる記入	置き換え後の文
もみじは見頃を迎え始めています	TREE は TREE-STATE を迎え始めています
松栄堂のタイムセールはお得です	SHOP の SALE はお得です
河原町のミュージアムは終了間近です	PLACE の EVENT は終了間近です

4.3 SC-LSTM を用いた観光案内文生成（提案手法）

前節で述べたコーパスをもとに文生成を行う．提案手法では，言語モデルとコンテンツベクトルの同時最適化を行う事で，テンプレートに依存しない文生成を行う．学習するモデルには 3.3.2 節で述べた SC-LSTM[37] を用いる．生成プロセスを図 4.1 に示す．

生成プロセスは，大きく文の生成フェーズとリランキングフェーズに分かれる．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 (CV) を入力として与え，SC-LSTM ジェネレータを用いて文候補を生成し，言語モデルによるエントロピ [7]・スロットエラーレートの計算を行う．リランキングフェーズでは，LSTM 言語モデルを用いて文候補のリランキングを行う．Wen ら

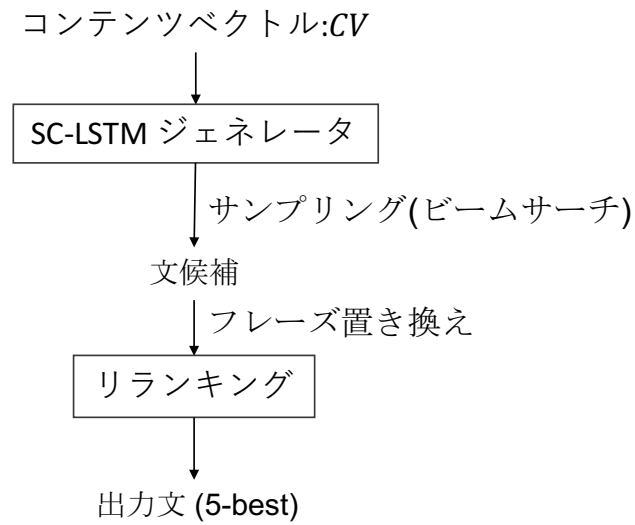


図 4.1 文生成プロセス（提案手法）

のアーキテクチャ [37] では、文生成を行った後に共通表現を含んだままリランキングを行っており、フレーズを考慮した文生成がなされない可能性がある。そこで本論文では、共通表現にフレーズを埋め込まない場合と埋め込む場合においてそれぞれ構築した LSTM 言語モデルを用いてリランキングを行い、生成文の比較を行う。リランキングコストとして、SC-LSTM 言語モデル（フレーズリランカ、共通表現リランカ）のエントロピー、LSTM 言語モデルのエントロピー、SERR を用いる。SERR はスロットエラーレートのことであり、コンテンツパターンと生成文に含まれるコンテンツにおける不足率・超過率を足し合わせたものである。この生成プロセスにより得られた上位 5 文を出力文とする。

4.4 テンプレートによる観光案内文生成（ベースライン）

ベースラインには，3.2 節で述べたテンプレートマッチングによる文生成を行う．生成プロセスを図 4.2 に示す．

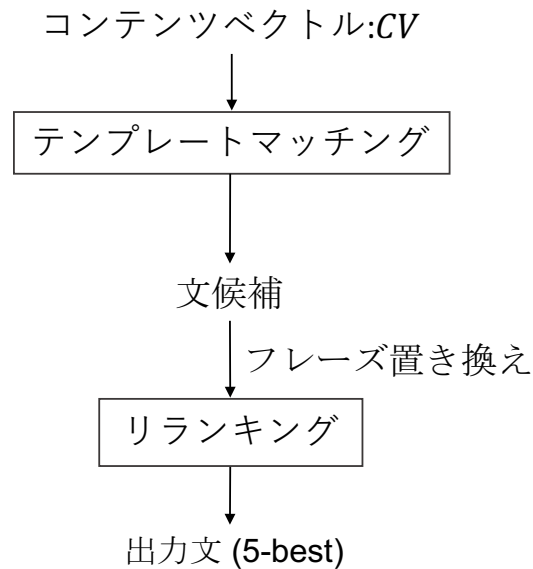


図 4.2 生成プロセス（ベースライン）

テンプレートマッチングによる文生成では，与えられたコンテンツパターン（コンテンツベクトル）に従ってテンプレートを抽出し，フレーズを埋め込む事で出力文となる．テンプレートとして，表 4.4 にある共通表現置き換え後の文をテンプレートとみなし，付随するコンテンツパターンとのマッチングを行う．具体的な生成プロセスは以下ようになる．

まず生成フェーズでは，与えられたコンテンツパターンとトレーニングコーパスにおけるコンテンツパターンを比較し，レーベンシュタイン距離の最も小さいコンテンツパターンを持つテンプレートを候補とする．その後，共通表現をフレーズに置き換えた後，LSTM 言語モデルによるエントロピーを用いてリランキングを行い，上位 5 文を出力文とする．レーベンシュタイン距離によるマッチングを行うため，この距離が 0 でない時にフレーズを埋め込むことが出来ない場合がある．その際は，対応するエンティティからランダムにフレーズを埋め込む．これは，LSTM 言

語モデルを用いてリランキングを行う際に、エンティティタグが含まれない状態にすることで適切なエントロピー計算を行わせるためである。この際に使用する LSTM 言語モデルは提案手法と同様のものである。

4.5 まとめ

本章では、観光案内コーパスを収集するにあたって必要となるコンテンツベクトルの設計について述べた。Wen らの提案した応答文生成システムでは、米国 SF サンフランシスコのレストランとホテルの対話を行うシステムにおいて用いられるエンティティ [38] に基いて設計されたコンテンツを使用している。こうしたエンティティに基づく設計に加え、本研究では、リアルタイムに変化する状況を説明可能とするような観光案内システムを実現するために「Begin」「Stay」「End」といった「始まったのか」「続いているのか」「終わったのか」を示すコンテンツを設定し、これらを用いて表現可能な 69 通りのコンテンツパターンを用いて、クラウドソーシングによるコーパス収集を行った。

観光案内文を生成する手法として、テンプレートを使用するもので近年提案されているテンプレートマッチングによる文生成と、SC-LSTM を用いた文生成を用いて、意図したコンテンツパターンに従った観光案内文生成が可能であるかどうかについて実験、評価を行う。

SC-LSTM を用いた文生成では、エンティティに埋め込むフレーズによる言い回しの変化を考慮できるよう、フレーズを埋め込んだ後に LSTM 言語モデルを用いてリランキングを行う手法を提案し、これによる観光案内文生成についても同様に実験、評価を行う。

第 5 章 評価実験

4 章にて述べたコーパス・文生成手法を用いて評価実験を行う。まず実験設定について述べ (5.1 節), 実験結果 (5.2 節), 最後に事例の分析を行う (5.3 節)。

5.1 実験設定

本システムでは, コンテンツベクトルを入力として観光案内説明文を生成する。本章では, 提案手法が有効であるかどうかを確認するため, 文生成を行い得られた説明文の評価を行う。

5.1.1 コーパスの前処理

言語モデルおよびコンテンツベクトルの学習を行うにあたって, 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 69 パターンのコンテンツベクトルを用いる。コンテンツベクトル 69 パターンのうち無作為に抽出した 10 パターンのコンテンツベクトルとそれに対応するアノテーション 1405 文を未知のベクトルパターンとして除外し, 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う。

学習を行うために, 前処理として, ワーカに指示したコンテンツにより得られたアノテーションを, 表 4.2 に従って共通表現に置き換え, コンテンツパターンに従ってコンテンツベクトルを作成する。その後 MeCab*を用いて分かち書きを行い, テンプレートの構築及び SC-LSTM ジェネレータ, LSTM 言語モデルの学習を行う。

*MeCab : <http://taku910.github.io/mecab/>

5.1.2 評価指標

文生成は、テストコーパスにおけるアノテーションを教師としたものと、予め除外したコンテンツパターンにおけるアノテーションを教師とした 2 種類について行い、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」と呼び、この 2 種類について評価する。

生成文の評価を行うにあたっては、テンプレートを用いた言語生成の先行研究 [28] に習い BLEU スコア [39](Bilingual Evaluation Undersyudy Score) と主観評価を用いる。BLEU は機械翻訳における翻訳文が参照訳とどれだけ近しいかを測るものであり、M 文の参照訳 $R = \{\{r_1^1, \dots, r_M^1\}, \dots, \{r_1^N, \dots, r_M^N\}\}$ と N 文の翻訳結果 $E = \{e^1, \dots, e^N\}$ の n-gram の一致する割合を下記により算出する。

$$BLEU(R, E) = \prod_{n=1}^4 \left(\frac{\sum_{i=1}^N m_n(r_1^i, \dots, r_M^i, e^i)}{\sum_{i=1}^N c_n(e^i)} \right) \cdot BP(R, E) \quad (5.1.1)$$

$$BP(R, E) = \min \left(1, \exp \left(1 - \frac{\sum_{i=1}^N |\tilde{r}^i|}{\sum_{i=1}^N |e^i|} \right) \right) \quad (5.1.2)$$

本研究における翻訳文は生成された観光案内文であり、参照訳はコンテンツベクトルを与えられたクラウドワーカが記入した文となる。

主観評価では、各手法において既知及び未知のベクトルパターンを入力とした文生成を 300 文程行い、自然性・妥当性・嗜好性により評価した。自然性では、生成文について「1. 違和感がある」～「5. 違和感はない」までの 1～5 までの 5 段階、妥当性では「1. 妥当ではない」～「5. 妥当である」までの 1～5 までの 5 段階の評価をクラウドワークスを用いて記入してもらう。嗜好性では、入力として用いたベクトルパターンにおいて生成された 3 つの手法のうちのそれぞれ 2 つについて、どちらが観光案内文として好ましいかを「-2. 前者である」～「2. 後者である」とした -2～+2 までの 5 段階評価として記入してもらう。評価者によるスコアの変動を考慮するため、各生成文における記入は 3 人のクラウドワーカによって行ってもらい、その平均をスコアとする。作業を行ってもらったワーカの数 は男性 23 人、女性 11 人の合計 34 名であり、年齢の平均及び標準偏差はそれぞれ 40.1, 10.8 であった。

5.2 結果

5.2.1 BLEU スコアによる比較

BLEU スコアにおける評価結果を表 5.1 に示す.

表 5.1 BLEU スコア

既知のコンテンツパターンによる文生成		
手法	ビーム幅	BLEU スコア
テンプレートマッチング	-	<u>47.70</u>
SCLSTM+ 共通表現リランカ	10	46.55
SCLSTM+ 共通表現リランカ	20	45.61
SCLSTM+ 共通表現リランカ	30	45.01
SCLSTM+ フレーズリランカ	10	46.72
SCLSTM+ フレーズリランカ	20	45.56
SCLSTM+ フレーズリランカ	30	44.89
未知のコンテンツパターンによる文生成		
テンプレートマッチング	-	37.99
SCLSTM+ 共通表現リランカ	10	<u>41.09</u>
SCLSTM+ 共通表現リランカ	20	38.88
SCLSTM+ 共通表現リランカ	30	38.90
SCLSTM+ フレーズリランカ	10	41.07
SCLSTM+ フレーズリランカ	20	38.90
SCLSTM+ フレーズリランカ	30	38.72

共通表現リランカは生成文に共通表現を含んだままリランキングを行うもの、フレーズリランカはフレーズ埋め込みを行った後にリランキングを行うものである。評価結果について、以下のような傾向が伺える。

- (i) 未知・既知のコンテンツパターンにおける BLEU スコアの違いについては、未知のコンテンツパターンによる BLEU スコアが既知のコンテンツパターンにおける BLEU スコアに比べて全手法が低下する結果となった。減少量

は、テンプレートマッチングについては大きく低下し、共通表現リランカとフレーズリランカにおいては、テンプレートマッチングに比べて減少量は少なくなった。

- (ii) 既知のコンテンツパターンによる文生成については、テンプレートマッチングが最も高くなった。共通表現リランカ、フレーズリランカについては、ビーム幅が 10 の場合において最も BLEU スコアが高く、全手法において BLEU スコアの大きな差異は見られなかった。
- (iii) 未知のコンテンツパターンにおける文生成については、テンプレートマッチングが最も低いスコアとなった。共通表現リランカ、フレーズリランカについては、ビーム幅が 10 の場合において最も BLEU スコアが高く、ビーム幅の違いによる BLEU スコアに差異が見られた。
- (iv) 共通表現リランカとフレーズリランカによる文生成については、BLEU スコアにおける同一ビーム幅同士の差異は見られなかった。

まず (i) についてであるが、未知のコンテンツパターンにおいてテンプレートマッチングの BLEU スコアが大きく低下した理由としては、生成の枠組み上、未知のコンテンツパターンに完全に対応した文生成を行うことができないためであるといえる。共通表現リランカとフレーズリランカについては減少量が少なく、未知のコンテンツパターンに対応した文生成が行われていると考察できる。

次に (ii) についてであるが、テンプレートマッチングによる手法で BLEU スコアが最も高くなった理由については、テンプレート方式はアノテーションを基にフレーズを埋め込む方式をとっており、ほぼアノテーションと相違の無い文生成を行う事が可能であるためではないかと考えられる。共通表現リランカとフレーズリランカによる BLEU スコアとの差は少なく、既知のコンテンツパターンにおいてはテンプレートマッチングには劣るものの、比較的近い文生成が行われているのではないかと推察できる。

次に (iii) についてであるが、テンプレートマッチングにおける BLEU スコアに比べて共通表現リランカ、フレーズリランカの BLEU スコアが向上していることから、共通表現リランカ、フレーズリランカにおいて未知のコンテンツパターンを考慮した文生成が行われているのではないかと考えられる。共通表現リランカ、フレーズリランカの両手法において、ビーム幅を広く取るほど BLEU スコアが低下

する結果となった。ビーム幅を広く取る場合、SC-LSTM ジェネレータの生成する文候補は多くなり、コンテンツベクトルによる制約に従わない文が含まれる可能性も高くなる。これにより、LSTM リランカがベクトル制約を受けない文を上位としてしまい BLEU スコアが低下したのではないかと考えられる。

最後に (iv) についてであるが、BLEU スコアによる比較においては、共通表現を含んだままリランキングを行っても、フレーズを考慮してリランキングした出力文と同程度の文生成が行われているという結果となった。この理由としては、クラウドワークスにおいてワーカの記入する文が、埋め込むフレーズによる影響を受けない共通的なものであったためではないかと思われる。例えば「公演会」であれば「開演している」、「ミュージアム」であれば「開かれている」といったフレーズによる接続の変化が考えられるが、両イベントにおいて「行われている」といった表現を行う事が可能であり、このような記入が多い場合においてはフレーズを埋め込まずとも違和感のない文を出力することができる。実際に観光案内において説明するコンテンツが増加するとともに接続を考慮する必要がある場面が増えていくと考えられ、その場合においてはフレーズを埋め込んだリランキングの方が適切な文を出力出来るのではないかと推察される。

5.3 主観評価

主観評価における結果を表 5.2, 表 5.3 に示す.

表 5.2 主観評価-自然性・妥当性

既知のコンテンツパターンによる文生成		
手法	自然性	妥当性
テンプレートマッチング	<u>4.01</u>	4.20
SCLSTM+ 共通表現リランカ	4.00	4.22
SCLSTM+ フレーズリランカ	3.96	<u>4.24</u>
未知のコンテンツパターンによる文生成		
テンプレートマッチング	3.98	3.10
SCLSTM+ 共通表現リランカ	<u>4.00</u>	<u>4.10</u>
SCLSTM+ フレーズリランカ	3.75	3.94

表 5.2 は, 生成文を「自然性」及び「妥当性」について「1~5」で 5 段階評価したものである. 表 5.3 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フレーズリランカによる手法が共通表現リランカによる手法に比べて劣る結果となった. フレーズリランカで用いられる 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.

表 5.3 主観評価-嗜好

既知のコンテンツパターンによる文生成			
Pref	テンプレート マッチング	SCLSTM 共通表現リランカ	SCLSTM フレーズリランカ
テンプレート マッチング	-	0.026	0.041
SCLSTM 共通表現リランカ	-0.026	-	0.013
SCLSTM フレーズリランカ	-0.041	-0.013	-
未知のコンテンツパターンによる文生成			
Pref	テンプレート マッチング	SCLSTM 共通表現リランカ	SCLSTM フレーズリランカ
テンプレート マッチング	-	-0.561	-0.332
SCLSTM 共通表現リランカ	0.561	-	0.232
SCLSTM フレーズリランカ	0.332	-0.232	-

5.4 事例分析

実際に生成された観光案内説明文について、生成例を用いて分析を行う。生成例は、テンプレートマッチング、共通表現リランカ、フレーズリランカの3種類について示す、共通表現リランカ及びフレーズリランカについては、BLEU スコアの高かったビーム幅が 10 の場合における生成例を示す。

5.4.1 既知のコンテンツベクトルによる文生成

生成例を表 5.4 に示す。既知のコンテンツパターンを用いた文生成においては、どの手法においても、コンテンツパターンの指示を満たした文生成が行われている事が多い。生成文の傾向としては、パターン (A) のような全く同じ文が生成される場合が多いが、パターン (B) のような意味は同じであるが異なる文生成も行われていた。共通表現リランカ、フレーズリランカの両手法において、テンプレートマッチングによる文生成と同程度の品質の文が出力された。

パターン (C) のような指示では、共通表現リランカを用いた文生成において「お特別展～」といった不自然な文生成が行われる例が見られた。これは「EVENT」エンティティに含まれている「祭り」といったようなフレーズをワーカが用いる際に、接頭語を付け加えて観光案内文の記入を行ったためであると考えられる。共通表現リランカではこのような言い回しに対応できない反面、フレーズリランカによる手法では接頭語を考慮したリランキングが行われ、文が出力されている。

表 5.4 既知のコンテンツベクトルによる生成結果

コンテンツパターン (A)	PLACE=上賀茂神社 EVENT=特別展 START
テンプレートマッチング	上賀茂神社 では 特別展 が 始まっ て い ます。
SCLSTM+ 共通表現リランカ	上賀茂神社 では 特別展 が 始まっ て い ます。
SCLSTM+ フレーズリランカ	上賀茂神社 では 特別展 が 始まっ て い ます。
コンテンツパターン (B)	SALE=限定販売 START
テンプレートマッチング	限定販売 が 始まろ う と し て い ます。
SCLSTM+ 共通表現リランカ	限定販売 が 始まっ て い ます。
SCLSTM+ フレーズリランカ	限定販売 が 始ま り ま し た。
コンテンツパターン (C)	EVENT=特別展 STAY
テンプレートマッチング	特別展 が 行わ れ て い ます。
SCLSTM+ 共通表現リランカ	お 特別展 が 続い て い ます。
SCLSTM+ フレーズリランカ	特別展 が 行わ れ て い ます。

5.4.2 未知のコンテンツベクトルによる文生成

表 5.5 未知のコンテンツベクトルによる生成結果

コンテンツパターン (A)	EVENT=祭り CONGESTION START
テンプレートマッチング	平等院 では 祭り が 行われ、 混雑 し はじめて います。
SCLSTM+ 共通表現リランカ	祭り で 混雑 が 始まっ て います。
SCLSTM+ フレーズリランカ	祭り は 混雑 し 始め て います。
コンテンツパターン (B)	SALE=割引セール GOODVALUE STAY
テンプレートマッチング	割引セール が 続い て います。
SCLSTM+ 共通表現リランカ	お 得 な 割引セール が 続い て います。
SCLSTM+ フレーズリランカ	お 得 な 割引セール が 続い て います。
コンテンツパターン (C)	PLACE=天満宮周辺 TREE=紅葉 BEAUTIFUL CONGESTION
テンプレートマッチング	天満宮周辺 では 美しい 紅葉 を 見よ う と する 人達 で 混雑 が 続い て います。
SCLSTM+ 共通表現リランカ	天満宮周辺 の 紅葉 が 綺麗 で 混雑 し て います。
SCLSTM+ フレーズリランカ	天満宮周辺 の 紅葉 が 綺麗 で 混雑 し て います。
コンテンツパターン (D)	PLACE=金閣寺 EVENT=お花見
テンプレートマッチング	金閣寺 では お花見 が 行われ て い て 混雑 し て います。
SCLSTM+ 共通表現リランカ	金閣寺 の お花見 は 混雑 し て います。
SCLSTM+ フレーズリランカ	金閣寺 の お花見 は 混雑 し て います。

生成例を表 5.5 に示す。テンプレートマッチングによる手法では、パターン (A)(B)(C)(D) 全てにおいて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」というフレーズが現れている。(B) では、「GOODVALUE」を指示しているにも関わらず、お買得といった表現がなされていない。(C)(D) では、「続いている～」や「混雑して～」といった、「STAY」「CONGESTION」等の指示を行っていないエンティティ情報を含んでいる。テンプレートマッチングによる手法では、指示したコンテンツパターンがコーパスに存在しない場合、最も指示コンテンツパターンに近いテンプレートを用いる事で生成を行っているため、未知のコンテンツパターンに対応することは難しい。

共通表現リランカ、フレーズリランカの両手法では、いくつかのパターンにおい

てはパターン (D) のように必要なコンテンツを述べられていなかったり、逆に不要なコンテンツを含んでいるような場合があった。しかしながら多くのコンテンツパターンにおいて、(A)(B)(C) のような、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事が確認された。

5.5 まとめ

本章では、4 章で述べた観光案内コーパスを用いて、テンプレートを用いた文生成と SC-LSTM を用いた文生成を行い、BLEU スコアによる比較及び生成例による分析を行った。

結果として、テンプレートマッチングでは対応することの難しい未知のコンテンツパターンを考慮した文生成が行われていることを確認し、観光案内システムにおける文生成について、言語モデルとコンテンツベクトルの同時学習を行う事で未知のコンテンツパターンを考慮した文が生成可能であることを事例により示した。

また、SC-LSTM を用いた文生成では、共通表現にフレーズを埋め込んだ後に LSTM 言語モデルを用いてリランキングを行う手法を提案し、提案手法において、接頭語によるフレーズ変化をふまえたリランキングが行われていることが確認できた。

第 6 章 結言

6.1 本論文のまとめ

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中間表現として定まっている時の文生成について検討を行った。従来の文生成では予め構築したテンプレートに中間表現を当てはめることで文生成を行うが、未知のエンティティの組み合わせを考慮した文生成が出来ず、中間表現がフレーズに限られるといった課題があった。本研究では言語モデルとコンテキストを同時学習させることで文生成を行う対話システムの先行研究のアーキテクチャをベースに、フレーズを埋めた後に LSTM 言語モデルを用いてリランキングを行う手法を提案し、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テンプレートマッチングによる手法より性能が向上させることを確認した。また、一部のパターンにおいて、先行研究で考慮されていなかった接頭語等を考慮した文生成が可能であることを確認した。

6.2 今後の課題

今後の課題として、以下が挙げられる。1 つ目に、コンテンツベクトルの設計をより具体化する事が挙げられる。今回想定する観光案内システムにおいては、実施したアンケートの解答を基に伝えたいコンテンツをいくつか固定して設計したが、実際にシステムに搭載するにあたっては、より広いドメインにおいて観光案内を行う必要がある。それに付随して、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させる必要がある。

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。現状のアーキテクチャでは、1 つのエンティティに対して 1 つのインデックスを割り当てているため、例えば「地名」を二度述べる事は難しい。また、コンテンツによっては生成する順番によって意味が変化する場合があり、本研究で提案観光案内システムを実用化させるためには、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

本学情報科学研究科の中村哲教授には、主指導教官として貴重なご助言を頂いたり、研究の方向性について多大なるご支援を賜りました。素晴らしい研究環境を与えて頂き、心より感謝を申し上げます。

本学情報科学研究科の松本裕治教授には、副指導教官として様々なご助言を賜り、心より感謝申し上げます。

本学情報科学研究科の吉野幸一郎助教には、いつも悩んでいた私に研究の方向性やご助言など数え切れない程のご支援を賜りました事、心より感謝を申し上げます。先生の支え無しにここまで来るのはあり得ませんでした。

本学情報科学研究科の鈴木優特任准教授には、研究のご支援初めクラウドソーシングやコーパス収集等においての貴重かつ的確なご助言を頂きました。心より感謝を申し上げます。

本学情報科学研究科の須藤克仁准教授，Sakriani Sakti 特任准教授，田中 宏季助教には、研究における貴重なご助言を頂きました。心より感謝申し上げます。

また、林美保様，松田真奈美様には研究以外の様々な活動や運営においてご尽力いただきました。心より感謝を申し上げます。

研究室の学生の皆様には、研究に対するご助言をいただいたり、励ましを頂きました。心より感謝を申し上げます。

最後に、私をここまで育て、たゆまぬ支援をしてくれた両親に心より感謝を申し上げます。

参考文献

- [1] D. Gavalas, C. Konstantopoulos, K. Mastakas, and G. Pantziou, “Mobile recommender systems in tourism,” *Journal of network and computer applications*, vol. 39, pp. 319–333, 2014.
- [2] R. Barta, C. Feilmayr, B. Pröll, C. Grün, and H. Werthner, “Covering the semantic space of tourism: an approach based on modularized ontologies,” in *Proceedings of the 1st Workshop on Context, Information and Ontologies*. ACM, 2009, p. 1.
- [3] M. Kenteris, D. Gavalas, and D. Economou, “An innovative mobile electronic tourist guide application,” *Personal and ubiquitous computing*, vol. 13, no. 2, pp. 103–118, 2009.
- [4] M. Rey-López, A. B. Barragáns-Martínez, A. Peleteiro, F. A. Mikic-Fonte, and J. C. Burguillo, “moretourism: Mobile recommendations for tourism,” in *Consumer Electronics (ICCE), 2011 IEEE International Conference on*. IEEE, 2011, pp. 347–348.
- [5] L. Sebastia, I. Garcia, E. Onaindia, and C. Guzman, “e-tourism: a tourist recommendation and planning application,” *International Journal on Artificial Intelligence Tools*, vol. 18, no. 05, pp. 717–738, 2009.
- [6] S. Venkataiah, N. Sharda, and M. Ponnada, *A Comparative Study of Continuous and Discrete Visualisation of Tourism Information*. Vienna: Springer Vienna, 2008, pp. 12–23. [Online]. Available: https://doi.org/10.1007/978-3-211-77280-5_2
- [7] 確率的言語モデル, ser. 言語と計算. 東京大学出版会, 1999. [Online]. Available: <https://books.google.co.jp/books?id=I78tAAAAMAAJ>
- [8] O. Averjanova, F. Ricci, and Q. N. Nguyen, “Map-based interaction with a conversational mobile recommender system,” in *Mobile Ubiquitous Computing, Systems, Services and Technologies, 2008. UBICOMM’08. The Second International Conference on*. IEEE, 2008, pp. 212–218.
- [9] J. M. Noguera, M. J. Barranco, R. J. Segura, and L. MartíNez, “A mobile

- 3d-gis hybrid recommender system for tourism,” *Information Sciences*, vol. 215, pp. 37–52, 2012.
- [10] K. Cheverst, N. Davies, K. Mitchell, A. Friday, and C. Efstratiou, “Developing a context-aware electronic tourist guide: some issues and experiences,” in *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. ACM, 2000, pp. 17–24.
 - [11] N. S. Savage, M. Baranski, N. E. Chavez, and y. Tobias Höllerer, “I’m feeling loco: A location based context aware recommendation system.”
 - [12] F. Ricci and Q. N. Nguyen, “Acquiring and revising preferences in a critique-based mobile recommender system,” *IEEE Intelligent systems*, vol. 22, no. 3, 2007.
 - [13] D. Martin, A. Alzua, and C. Lamsfus, “A contextual geofencing mobile tourism service,” in *ENTER*, 2011, pp. 191–202.
 - [14] G. Tumas and F. Ricci, “Personalized mobile city transport advisory system,” *Information and Communication Technologies in Tourism 2009*, pp. 173–183, 2009.
 - [15] B. Zenker and B. Ludwig, “Rose: assisting pedestrians to find preferred events and comfortable public transport connections,” in *Proceedings of the 6th International Conference on Mobile Technology, Application & Systems*. ACM, 2009, p. 16.
 - [16] A. Umanets, A. Ferreira, and N. Leite, “Guideme – a tourist guide with a recommender system and social interaction,” *Procedia Technology*, vol. 17, no. Supplement C, pp. 407 – 414, 2014, conference on Electronics, Telecommunications and Computers – CETC 2013. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2212017314004848>
 - [17] G. Abowd, A. Dey, P. Brown, N. Davies, M. Smith, and P. Steggles, “Towards a better understanding of context and context-awareness,” in *Handheld and ubiquitous computing*. Springer, 1999, pp. 304–307.
 - [18] C. Lamsfus, A. Alzua-Sorzabal, D. Martín, Z. Salvador, and A. Usandizaga, “Human-centric ontology-based context modelling in tourism.” in *KEOD*,

- 2009, pp. 424–434.
- [19] J. M. Conroy and D. P. O’leary, “Text summarization via hidden markov models,” in *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval*. ACM, 2001, pp. 406–407.
 - [20] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
 - [21] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” in *Advances in neural information processing systems*, 2014, pp. 3104–3112.
 - [22] M. Agirrezabal, B. Arrieta, A. Astigarraga, and M. Hulden, “Pos-tag based poetry generation with wordnet,” in *ENLG*, 2013, pp. 162–166.
 - [23] X. Zhang and M. Lapata, “Chinese poetry generation with recurrent neural networks,” in *EMNLP*, 2014, pp. 670–680.
 - [24] H. Horacek, *The Architecture of a Generation Component in a Complete Natural Language System*. Academic Pres, 1990.
 - [25] R. Dale, S. Geldof, and J.-P. Prost, “Coral: Using natural language generation for navigational assistance,” in *Proceedings of the 26th Australasian Computer Science Conference - Volume 16*, ser. ACSC ’03. Darlinghurst, Australia, Australia: Australian Computer Society, Inc., 2003, pp. 35–44. [Online]. Available: <http://dl.acm.org/citation.cfm?id=783106.783111>
 - [26] S. Sripada, E. Reiter, and I. Davy, “Sumtime-mousam: Configurable marine weather forecast generator,” *Expert Update*, vol. 6, no. 3, pp. 4–10, 2003.
 - [27] 亀甲博貴, 三輪誠, 鶴岡慶雅, 森信介, 近山隆 *et al.*, “対数線形言語モデルを用いた将棋解説文の自動生成,” *情報処理学会論文誌*, vol. 55, no. 11, pp. 2431–2440, 2014.

- [28] 山崎健史, 吉野幸一郎, 前田浩邦, 笹田鉄郎, 橋本敦史, 船富卓哉, 山肩洋子, 森信介 *et al.*, “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.
- [29] W. C. Mann and S. A. Thompson, “Rhetorical structure theory: Toward a functional theory of text organization,” *Text-Interdisciplinary Journal for the Study of Discourse*, vol. 8, no. 3, pp. 243–281, 1988.
- [30] A. Stent and M. Molina, “Evaluating automatic extraction of rules for sentence plan construction,” in *Proceedings of the SIGDIAL 2009 Conference: The 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, ser. SIGDIAL ’09. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009, pp. 290–297. [Online]. Available: <http://dl.acm.org/citation.cfm?id=1708376.1708417>
- [31] A. Ratnaparkhi, “Trainable approaches to surface natural language generation and their application to conversational dialog systems,” *Computer Speech & Language*, vol. 16, no. 3, pp. 435–455, 2002.
- [32] R. Kondadadi, B. Howald, and F. Schilder, “A statistical nlg framework for aggregated planning and realization.” in *ACL (1)*, 2013, pp. 1406–1415.
- [33] G. Angeli, P. Liang, and D. Klein, “A simple domain-independent probabilistic approach to generation,” in *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 2010, pp. 502–512.
- [34] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, “Recurrent neural network based language model.” in *Interspeech*, vol. 2, 2010, p. 3.
- [35] S. Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 6, no. 02, pp. 107–116, 1998.
- [36] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [37] T.-H. Wen, M. Gasic, N. Mrksic, P.-H. Su, D. Vandyke, and S. Young, “Se-

- mentally conditioned lstm-based natural language generation for spoken dialogue systems,” *arXiv preprint arXiv:1508.01745*, 2015.
- [38] M. Gašić, D. Kim, P. Tsiakoulis, and S. Young, “Distributed dialogue policies for multi-domain statistical dialogue management,” in *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 5371–5375.
- [39] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “Bleu: a method for automatic evaluation of machine translation,” in *Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics, 2002, pp. 311–318.

付録

A 実施アンケート及び得られた解答

Q1. 観測データからどの様なデータを取得できる/できそうですか？

A1.

ユーザの加速度，歩行速度・加速度，位置情報，目的地，気温や降水量，次行動
音声，各観光スポットに関する写真やコメント，観光客の報告情報
各観光スポットの混雑度，ユーザプロフィール 等

Q2. 取得されたデータから，どのようなフレーズが得られる/得られそうですか？

A2.

観光地名，店名，神社，桜の状態，イベント情報，名物，交通量
混雑状態，綺麗かどうか，おいしい

Q3. どのような観光案内を行う事ができる/できそうですか？

A3.

土産物・食品などを買うお薦めの店，お薦めのレストラン
観光地の混雑度や天気 紅葉などの様子 飲食店の混雑度
桜を撮るのに今のお薦めスポット、初詣などのイベント時の今の混雑状況や天気
途中の変更にうまく対応してリスケしてくれる案内システム
おすすめのスポット（混雑度を考慮した）の提案 等

Q3. どの様なシーンで，どんな観光案内をして貰えると有用だと思いますか？

A3.

混雑しないルートでの観光案内を行うシステム
京都の評判が良い・悪いスポットやイベント情報
人数に合わせたスポットをリコメンドする観光案内
自分の趣味嗜好やリアルタイムに変化する情報に合わせて観光スポットを推薦
その場になくても遠くの施設の現在の状況を教えてくれる（混雑状況、画像など）案内
今から行こうと思っている観光地などの天気（急な天気の変化）
現在地付近の飲食店情報（混雑度やその日が定休日かどうかなど） 等

B 設定したコンテンツパターン

ベクトルパターン		
Tree,beautiful, Tree,beautiful,congestion, Tree,Tree-state, Tree,Tree-state,beautiful, Tree,Tree-state,congestion, place,Tree,beautiful, place,Tree,beautiful,congestion, place,Tree,Tree-state, place,Tree,Tree-state,beautiful, place,Tree,Tree-state,congestion, event,begin, event,congestion, event,congestion,end place,event,stay, place,event,congestion,begin, sale,begin, shop,sale, shop,sale,end sale,Good-value,stay, shop,sale,Good-value,stay, shop,sale,congestion,stay, shop,congestion,begin, place,congestion,	Tree,beautiful,begin, Tree,beautiful,congestion,begin, Tree,Tree-state,begin, Tree,Tree-state,beautiful,begin, Tree,Tree-state,congestion,begin, place,Tree,beautiful,begin, place,Tree,beautiful,congestion,begin, place,Tree,Tree-state,begin, place,Tree,Tree-state,beautiful,begin, place,Tree,Tree-state,congestion,begin, event,stay, event,congestion,begin, place,event, place,event,end place,event,congestion,stay, sale,stay, place,shop,sale,begin, sale,Good-value, shop,sale,Good-value, place,shop,sale,congestion, shop,sale,congestion,end shop,congestion,stay, place,congestion,begin,	Tree,beautiful,stay, Tree,beautiful,congestion,stay, Tree,Tree-state,stay, Tree,Tree-state,beautiful,stay, Tree,Tree-state,congestion,stay, place,Tree,beautiful,stay, place,Tree,beautiful,congestion,stay, place,Tree,Tree-state,stay, place,Tree,Tree-state,beautiful,stay, place,Tree,Tree-state,congestion,stay, event,end event,congestion,stay, place,event,begin, place,event,congestion, place,event,congestion,end sale,end shop,sale,stay, sale,Good-value,begin, place,shop,sale,Good-value,begin, shop,sale,congestion,begin, shop,congestion, shop,congestion,end place,congestion,stay,

発表リスト

- [1] 伝達内容を考慮して言語化するニューラル言語生成の検討, 研究報告自然言語処理 (NL), No.13, pp.1-5, 2017