

フローグラフからの手順書の生成

山崎 健史^{1,a)} 吉野 幸一郎² 前田 浩邦^{3,*} 笹田 鉄郎⁴
橋本 敦史⁵ 船富 卓哉² 山肩 洋子⁶ 森 信介^{4,b)}

受付日 2015年6月30日, 採録日 2015年12月7日

概要: コンピュータから人間へ情報を伝達するためには、コンピュータが処理しやすい表現を、人間に理解しやすい自然言語文に変換する技術が重要である。本研究では、構造化された情報であるフローグラフから自然言語文を自動生成する手法を提案する。手順書のドメインとしてレシピを扱う。まず、実際の手順書に対して重要語とその種類を表すラベルを自動認識し、固有表現に置き換えることで、統計的なテンプレートの自動生成を行う。次に、フローグラフを1つの動作を含む最小手順部分グラフに分割する。最後に、最小手順部分グラフに最もマッチするテンプレートを選択して文生成を行う。テンプレート選択の指標として、負対数確率、内容語残存コスト、エントロピーを用いる。実験では、正解として用意されたフローグラフを表現する自然言語文章と生成された文章を比較評価した。具体的にはBLEU値による自動評価と主観評価を行い、提案手法が精度向上に寄与し、BLEUが人手による主観評価と相関があることを確認した。

キーワード: 自然言語生成, 手順書, テンプレート

Procedural Text Generation from a Flow Graph

TAKESHI YAMASAKI^{1,a)} KOICHIRO YOSHINO² HIROKUNI MAETA^{3,*} TETSURO SASADA⁴
ATSUSHI HASHIMOTO⁵ TAKUYA FUNATOMI² YOKO YAMAKATA⁶ SHINSUKE MORI^{4,b)}

Received: June 30, 2015, Accepted: December 7, 2015

Abstract: Algorithms that convert a computational representation to human-readable natural language sentences are essential for communications. In this paper, we propose a method that automatically generates natural language sentences from a structured flow graph. We demonstrate this method in the domain of cooking recipes. Firstly, we automatically generate templates from real recipe texts by recognizing keywords and their type label, and replacing the words with named entity labels. Secondly, we divide the flow graph into minimum procedural subgraphs, each of which has one action. Finally, we select the most suitable template for each subgraph. When selecting a template, we use three measures: a negative logarithmic probability, a remaining content word cost, and a sentence entropy. In the experiments we generate recipe texts from flow graphs and evaluated them through both an objective metric BLEU and a subjective evaluation by human examines. The results showed that our method improves both metrics and BLEU is correlated to the subjective metrics.

Keywords: natural language generation, procedural text, template

¹ 京都大学大学院情報学研究科
Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
² 奈良先端科学技術大学院大学情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology, Ikoma, Nara 630-0192, Japan
³ サイボウズ株式会社
Cybozu, Inc., Chuo, Tokyo 103-6028, Japan
⁴ 京都大学学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University, Kyoto 606-8501, Japan

⁵ 京都大学大学院教育学研究科
Graduate School of Education, Kyoto University, Kyoto 606-8501, Japan
⁶ 東京大学大学院情報理工学系研究科
Graduate School of Information Science and Technology, The University of Tokyo, Bukyo, Tokyo 113-8656, Japan
a) yamasaki@ar.media.kyoto-u.ac.jp
b) mori@ar.media.kyoto-u.ac.jp
*本研究への貢献は、京都大学大学院情報学研究科修士課程に在籍中になされた。

1. はじめに

言語は人間にとって最も自然に扱えるインタフェースの1つである。コンピュータが日々の生活に浸透していくとともに、コンピュータから人間への情報伝達の一手段として自然言語生成技術に対する需要が高まってきた。たとえば、コンピュータが手順書を理解し、理解した内容と実施している作業の状況を認識するシステムを組み合わせることを想定したとき、言語による提示は、画面表示や音声といった作業中でも利用可能な伝達手段によって手順を伝えることができるため有用である。しかし、コンピュータが扱うことに最適化された表現を人間が一目で把握することは難しい。そこで、どのようにしてコンピュータが持つ情報を人間に伝えるかということが問題となる。本研究では、コンピュータが持つ情報が定まっている中で、どのような表現で言語生成を行うかに主眼を置き、こうした意味表現から音声での伝達が可能な表現への変換として、自然言語文の生成（以下、文生成と呼ぶ）を扱う。

本研究では、何らかの処理を行うために一連の手続きが記述されている文書を手順書とし、これを生成の対象とする。コンピュータが持つ情報として、フローグラフで表現された手順書の意味表現 [1], [2] を想定して文生成を行う。フローグラフは物体や動作に対応する節点とその関係を示す枝によって手順を表現し、映像の理解結果など、様々なものから生成されることを想定している。手順書を対象とする理由として2点あげられる。1つ目は、手順書は文書の中でも存在する節点や枝に曖昧性が少なく、フローグラフなどの意味表現で扱いやすい点である。2つ目は、フローグラフは物体（中間生成物を含む）と動作の半順序関係を表すため、文の順序が重要である手順書に適する点である。本論文における生成文の順序関係の評価は、6.4節の主観評価で間接的に考慮している。フローグラフの応用を想定した場合、モデルの学習に用いるテキストとモデルの適用対象のドメインは一致することが望ましい [3]。そこで本研究では、テンプレート生成に用いるコーパスやフローグラフでの表現をレシピドメインに限定して議論する。ここでレシピとは、ある特定の料理を完成させるために必要な一連の手続きを表す。様々な手順書が存在する中でレシピドメインを選択した理由として、大量のコーパスが容易に入手できる点と、我々の調査によれば、他の多くの手順書ドメインよりも表現がより多様で、かつ、語彙サイズが比較的大きいことから、レシピドメインで利用可能な手法を構築できれば、他のドメインへの応用が容易であることが想定される点があげられる。

フローグラフからの手順文の生成の応用として、以下の例があげられる。

映像認識からの文生成 映像 (Computer Vision) からの認識結果を意味表現として部分フローグラフに変換し、

これを入力として文生成を行う。映像から意味表現を得る研究として、調理映像の認識結果 [4], [5], [6], 映像認識研究としては調理や理科実験の作業映像に対する単独の物体や動作の認識 [7], [8] をはじめとして、動作と対象を同時に認識し述語項構造を出力する研究 [9] などがある。

フローグラフ間の演算 複数のフローグラフに対して演算を行うことが提案されている。演算結果の例として、同じ料理名の複数のレシピのフローグラフから計算される典型的な手順のフローグラフ [10] などがあげられる。これを入力とすれば、ある料理名の典型的な作り方のレシピを生成することができる。

調理者への作業の教示 コンピュータが調理者に次の動作などを教示する補助システムが考えられている (Smart Kitchen [11])。あらかじめ料理のフローグラフを記憶しておき、調理の進行状況と調理者の振舞いに応じて、次の動作を部分グラフから生成した自然言語文で教示する。

フローグラフ節点の翻訳による手順書の生成 フローグラフの各接点を表す単語列を別の言語に翻訳し、その結果得られるフローグラフから手順書を生成することで、その言語で書かれた手順書を生成する。

本研究では、大量のコーパスからテンプレートを作成しておき、コンピュータに最適化された情報であるフローグラフを入力として統計的な文生成を行う手法を提案する。こうした学習コーパスからテンプレートを学習する枠組みは、少量の対話データに対して行われてきた [12], [13], [14] が、本研究では発話ではなく手順書からテンプレート学習を行い文生成を行う。提案手法では、より大規模な生テキストの自動解析結果を用いて自動的にテンプレートプールを作成する。提案手法の設計と概要については2章で、フローグラフについては3章で述べる。また、提案手法の詳細については4章と5章で述べる。実験により、学習データ量が生成される文章の質に大きく影響することを示し、実際にそれを検証した。

本システムによって自動生成された文の評価を行う際に、意味的な正しさを自動的に評価することは困難であり、大規模な評価を行うことが難しい。そのため、本研究では本来は機械翻訳の評価に使われていた BLEU [15] が生成文の評価にも適用可能であるかどうかについても検討を行った。BLEU による自動評価と人手による主観評価の両方で本手法の評価を行い、自動評価と主観評価の相関を調査した。

手順書を生成するにあたっては、どのように表現するかの前に、そもそも話す内容を意味表現として作成する必要がある。本研究では、先行研究 [2], [10] により、話す内容の意味表現が得られていることを前提とする。本研究の貢献は、文献 [2], [10] から得られた意味表現の自然言語への変換であり、より詳細には次の (1)~(3) である。本研究

の貢献は以下のとおりである。

- (1) 大規模な自動テンプレートの自動学習・選択による手順書の生成。
 - (a) テンプレートを獲得する手法の提案 (4.2 節)
 - (b) テンプレートを選択する手法の提案 (4.3 節)
 - (c) 活用語に対する活用形の言語モデルによる自動選択 (5.3 節)
- (2) (1) が実験的にうまくいくことを示した (6.5 節)。
- (3) 本タスクにおいて BLEU による自動評価と人による生成文の理解に相関があることを示した (6.4 節)。

2. 手順書のための文生成システムの設計

本章では、手順書を自動生成するための文生成システムの設計について述べる。手順書においていうことは、材料や道具などの物理実体（主に名詞句）、その様態の表現（主に形容詞句）、さらに何をするか（動詞句）である。大量の実例に対してこれらを自動認識することで、いうことをその種類を表すクラスに置き換え、いい方だけを含む文骨格を自動的に大量に生成する。この文骨格は、人手で記述される場合にテンプレートと呼ばれており、本研究でもテンプレートと呼ぶ。文生成における課題は、いうことの自動認識における誤りの影響をどう低減させるか、どの程度の量の実例があれば本研究で提案する文生成システムが機能するのかであり、本論文の後半 6.3 節、6.5 節でこれらの疑問に答える。本システムの概要を図 1 に示す。以下では、図 1 の各部分について概説する。

2.1 テンプレート生成

本手法では、あらかじめ収集したコーパスから自動的にテンプレートプールを作成する。収集したレシピコーパスに対して自動重要語認識を行い、その部分を重要語のクラスを表す重要語タグで置き換える。この際、活用語の活用語尾をテンプレートに含めることを提案する。テンプレート生成については 4.2 節で述べる。一般に、テンプレートには活用語尾を含めず、テンプレートに活用語を埋めるときに適切な活用語尾を推定し付加する設計が行われてきた。この設計は、異なる活用形からなるテンプレートを統合することができるため、必要となるテンプレート数を大幅に

削減できる。しかし、適切な活用語尾の推定 [16] を行う必要があることから、特に未知語に対して脆弱であり、ドメイン依存性の高い手順書に適さないと考えられる。提案手法では、言語ごとに異なる特有の知見を用いることを避けられるため、活用形を生成する規則の記述が不要で、未知語の活用型を推定する必要がない。そのため、語幹に文字を接続することで活用形の文字列が得られる言語に本システムを適用することは容易であると考えられる。短所として、テンプレート数が多くなり、活用語の語幹と活用語尾が適切に接続できなくなる可能性があるが、本研究ではこの問題に対して、学習データ量に対するテンプレートのカバレッジを 6.3 節で、活用語尾の接続の精度を 6.6 節で提示し、問題が起こらないことを実験的に示す。

2.2 最小手順部分グラフ抽出

本研究では、入力に意味構造を表すフローグラフを想定している。文書の大域構造については、1 文で 1 つの指示とし、手続きの実施順に文とすることで十分であろう。なぜなら、手順書における文には、1 からただか数個の手続きしか含まれておらず、また、手続きどうしの関係が明快だからである。そこで、手順書を自然言語として生成するために、入力されたフローグラフを実施者を主語とする 1 つの動作（述語）を含む最小手順部分グラフに分割する。これは、手順書を 1 つ 1 つの手順へと分割することを想定した場合、動作を基準とした単位が意味的に最小となるからである。フローグラフについては 3 章、最小手順部分グラフ抽出については 4.1 節で詳しく述べる。

2.3 テンプレート選択

フローグラフから分割した最小手順部分グラフの中に含まれる重要語タグから、重要語タグ列を作成する。この重要語タグ列が使われるテンプレートを選択候補にあげる。本研究で提案する文生成システムでは、負対数確率（テンプレートの出現頻度）、内容語残存コスト、エントロピーの 3 指標を統合したテンプレート選択コストを用いて最適なテンプレートを選択する。テンプレート選択については 4.3 節で、テンプレート選択コストについては 5 章で詳しく述べる。

2.4 文生成

選択したテンプレートに、入力の最小手順部分グラフの重要語を代入し、手順文を生成する。文生成については 4.4 節で述べる。また、生成した手順文の評価を 6.4–6.6 節で行う。

3. フローグラフ

本研究で提案するシステムでは、入力としてフローグラフを用いる。フローグラフは物体や動作に対応する節点と

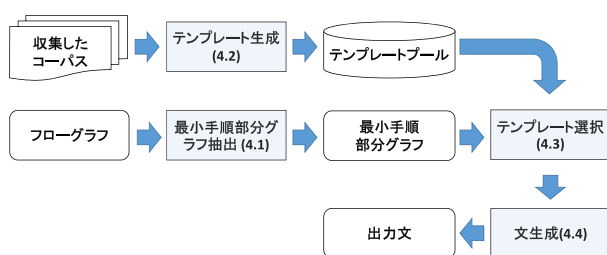


図 1 本システムの概要図

Fig. 1 Summary of our system.

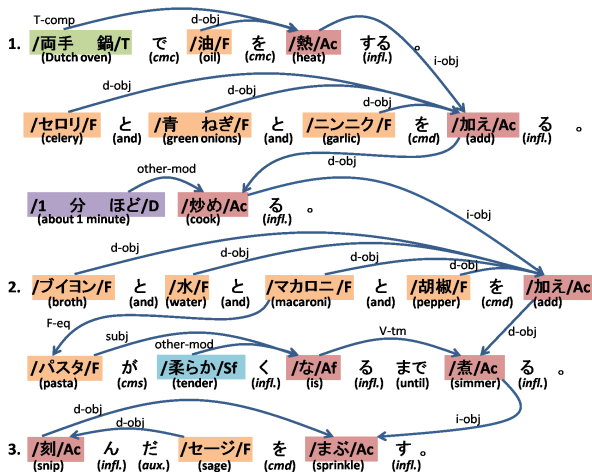


図 2 レシピテキストのフローグラフの例
Fig. 2 The flow graph of the example recipe.

表 1 レシピ固有表現の種類
Table 1 Types of r-NE.

タグ	意味	備考
F	食材	中間・最終生成物を含む
T	道具	調理道具や器など
D	継続時間	概数表現を含む
Q	分量	概数表現を含む
Ac	調理者の動作	語幹のみ
Af	食材の動作	語幹のみ
Sf	食材の状態	
St	道具の状態	

その関係を示す枝によって手順を表現するものであり、本研究ではドメインをレシピに限定したレシピフローグラフ [2] を用いる。レシピフローグラフとは、レシピの意味構造を単語ごとの関係で表した無閉路有向グラフである (図 2)。また、枝ごとに順序がない無順序グラフである。フローグラフの各節点のラベルは、単語列とレシピ用語タグ (r-NE タグ [2]) の組である。r-NE タグは調理を行ううえでの各単語列の役割を表しており、食材や道具などの名詞句や調理者の動作といった用言などに付与されている。フローグラフの各枝には述語項構造や共参照関係を表すための r-NE の関係を表すラベルが付与されている。しかし、1 章のフローグラフの応用例 (映像認識からの文生成、フローグラフ間の演算 [10]) では枝のラベルを用いることが想定されていないため、本研究では単語どうしの親子関係のみを用い、枝のラベルは用いない。表 1 に r-NE タグの一覧を示す。

フローグラフの節点どうしは有向枝によって接続されており、作業の依存関係を表す。たとえば、図 2 中で、“/加え/Ac” 節点は、“/熱/Ac” の作業の結果と“/セロリ/F”、“/青ねぎ/F”、“/ニンニク/F” を前提とした作業であることを表す。調理者の動作 (Ac) 間の有向枝は、その性質上ほぼ時間的順序を表すが、複数の作業者がいれば同時進

行が可能な場合も存在する。

以下では、各 r-NE タグについて順に簡単に説明する。

F: 食材 食材は、材料リストに記載されているような材料名に加えて、「タネ」「生地」といった中間食材名や、料理の完成品が含まれる。食せる、捨てる、量に変化するならば F と判定される。ただし修飾句は F に含まれない。

T: 道具 鍋、蓋、包丁、コンロなど、調理道具や器を道具としている。手や指などの調理者の身体の一部も道具のような振舞いをする場合は、T と判定される。食せない、量が変わらない点以外は、おおむね「食材」の判定と同じである。

D: 継続時間 加熱時間や冷却時間など、加工の継続時間を示している。数字と単位のほか、それらに対する概数表現なども含まれる。

Q: 分量 食材の量を指定する表現を分量としている。数字部分や単位に加えて、それらに対する概数表現なども含まれる。

Ac: 調理者の動作 調理者が主語となる用言で、主に他動詞である。文生成時の活用形の選択をテンプレート選択によって実現する設計のため、活用語尾は含まない。また、補助的動詞は含まれない。

Af: 食材の動作 食材が主語となる用言は、食材の動作としている。Ac と同じく、活用語尾は含まない。

Sf: 食材の様態 用意された食材様態の初期状態を表す表現、ならびに Ac や Af の影響で遷移する (した) 状態を表す表現を食材の様態としている。

St: 道具の様態 用意された道具様態の初期状態を表す表現、ならびに Ac や Af の影響で遷移する (した) 状態を表す表現を道具の様態としている。

4. 手順書の生成システム

本章では、本研究で提案する手順書生成システムについて述べる。本システムはフローグラフを入力としてレシピ文章を生成するシステムである。

本研究が扱う問題の難しさは、r-NE に接続される単語の自動選択にある。たとえば、T: 道具の後にくる単語の分布の上位は、に: 41.1%、で: 31.0%、を: 13.8%であった。この分布から、最頻値をとった場合でも 6 割は誤ってしまうため、決定的な手法は不適切であると考えられる。

まず 4.1 節でシステムに与える入力について説明し、その後システムの動作内容について述べる。本システムは大きく 3 つの工程に分かれており、(1) テンプレート生成 (4.2 節)、(2) テンプレート選択 (4.3 節)、(3) 文生成 (4.4 節) からなる。

4.1 入力: 最小手順部分グラフ抽出

本システムに与える入力は、レシピドメインの意味表現

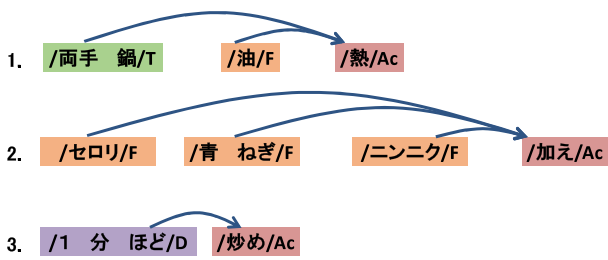


図3 フローグラフの最小手順部分グラフによる分割例

Fig. 3 Division of the example flow graph with minimum procedural subgraph.

であるフローグラフのAc（調理者の動作）節点を基準に分割した最小手順部分グラフ（以下、単に部分グラフと呼ぶ）である。これは、自然言語における意味表現を考える際に、動作（述語）を基準とする単位が意味的に最小となるからである。また、映像認識の出力結果を利用するような場合でも、調理者の個々の動作を基準に出力が分割されることが多い。以下の手順を行うことで、フローグラフから本システムの入力形式を得る。

- (1) Ac 節点を探索し、Ac 節点のリストを得る。
- (2) Ac 節点の出現順に、各 Ac 節点に対して、Ac 節点が根節点となるように部分グラフを構築する。このとき、1度使用した節点は再び使用しないようにする。

図2のフローグラフにおける第1手順から部分グラフによって分割グラフを抽出した例を図3に示す。

なお、元のフローグラフコーパスは、作成する段階で元のレシピに出現するNEの順序で作成されている。そのため、Ac 節点を出現順に探索して部分グラフの構築を行う本手法では、フローグラフにない情報（元のレシピのNEの順序）を用いて文生成を行っているといえる。一方、フローグラフのNEの順序が元のレシピの順序ではなかったとしても、フローグラフの作業順序を整理し、あらかじめ並べ替えを行うことで、本手法を追随することができる。このとき、並列の関係にある作業どうしの順序は、元のレシピのNEの順序が保証されないため、評価実験に用いるBLEU [15] の値は本手法に比較して低下する可能性がある。

本研究では述語間の関係を含むラベルの情報を用いおらず、順接やそれ以外の関係のラベルはフローグラフでは定義されていない [2], [10]。順接以外の述語間の関係の例として、

- 選択（例：ミョウガ/F は 小 さ/Sf く 輪切り/Ac もしくは みじん切り/Ac する）
- 条件（例：ワントン/F の 皮/F を 入れ/Ac する なら この時に）

などがあげられる（詳細は文献 [10] を参照）。本研究で用いているフローグラフにおいて、上記の例に関してはすべてのAcをフローグラフに含めて、すべて順接の関係としている。この例の場合は、実際に行わない作業が生成された文に含まれる。この問題を回避するために、このような述

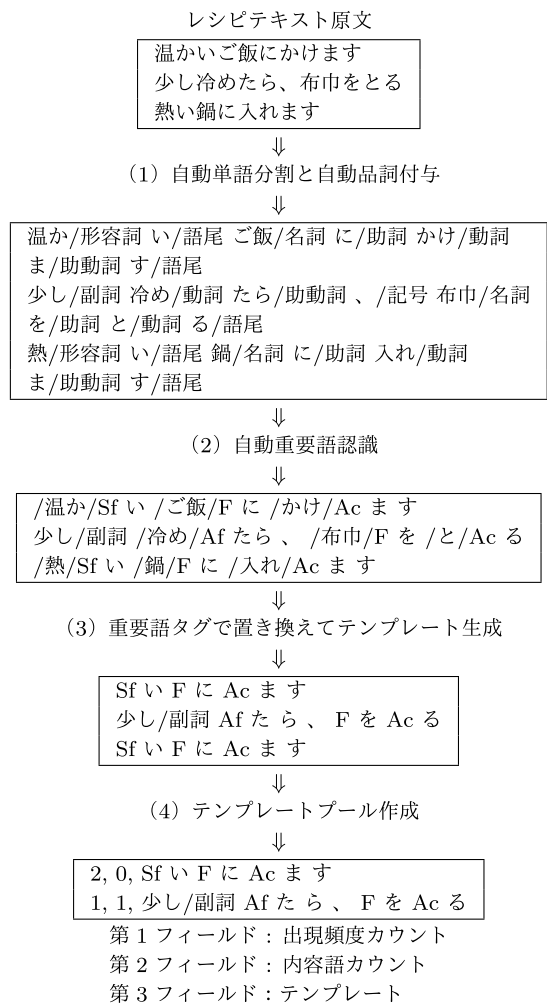


図4 テンプレート生成

Fig. 4 Generation of templates.

語間の関係を新しくフローグラフにラベルとしてアノテーションするなどの対応が考えられる。

4.2 テンプレート生成

本節では収集した手順書からテンプレートプールを作成する手法について述べる。テンプレートプールとは、r-NEの単語列をr-NEタグで置き換えた文骨格（テンプレート）の集合である。本手法では、テンプレート生成部ではr-NEタグに該当しない内容語を含むテンプレートを除去せず、テンプレート選択時に内容語を含むテンプレートに対してコストを設けて、選択される可能性を下げている。このような手法を選択する理由として以下の2点があげられる。

- 内容語の判別用いる品詞推定は自動推定のため、不要な内容語を誤りなく特定できない。
- テンプレート除去によるカバレッジの低下を避ける。

以下の手順で、レシピ文章を学習コーパスとして与え、自動的にテンプレートに変換する。また、一連の流れの具体例を図4に示す。

- (1) レシピ文章をテキスト解析ツール KyTea^{*1} [17] を用

*1 <http://www.phontron.com/KyTea/index-ja.html>

いて単語単位に分割し、品詞を付与する。

- (2) 単語列に対して r-NE 認識器：PWNER^{*2} [18] を用いて r-NE タグを付与する。
- (3) r-NE を r-NE タグで置き換える。
- (4) 同じテンプレートを統合し、テンプレートの出現頻度と、(1) の品詞情報をもとにテンプレートの内容語（後述：5.2 節）の個数をテンプレートの付随情報とする。テンプレートの出現頻度、および r-NE タグで置き換えられない内容語の個数を併記する。これらの情報は、テンプレート選択 (4.3 節) で用いる。なお、各テンプレート t_i はテンプレートのキー $k(t_i)$ で参照できるようにする。テンプレートのキー $k(t_i)$ は、 t_i 中に出現する r-NE タグを並べたものである。たとえば、図 4(4) の第 1 文の $k(t_i)$ は “Sf-F-Ac” となる。

4.3 テンプレート選択

4.2 節で作成したテンプレート群の中には、入力となる同一のフローグラフに対応するテンプレートが複数存在する。また、表現の適切さや一般性などが与えられておらず、このままでは一意にテンプレートを決定し文生成を行うことができない。そこで本研究では各テンプレートに対するコストを定義し、このコストによって適切なテンプレートを選択する。コストの定義の詳細については 5 章で述べる。テンプレート選択では、以下の処理でテンプレートを選択する。

- (1) 入力の部分グラフから、キーを生成する。フローグラフは無順序木なので、図 5 のように木構造の入替えが可能な場合、キー群として得る。
- (2) 各キーに対して、キーにマッチするテンプレート群を得て、各テンプレートのコストを計算する (5 章)。
- (3) (2) で計算したコストが最小となるテンプレートを選

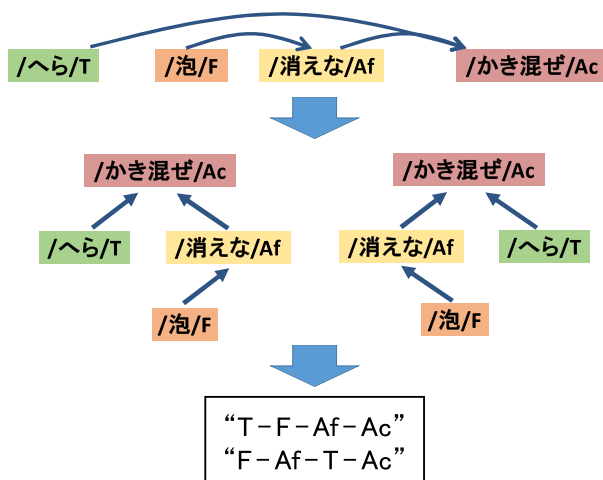


図 5 r-NE 列群の生成

Fig. 5 Generation of a group of r-NE sequence.

択する。

なお、(2) の手順は、最初にテンプレートを発見できたキーのみ処理を行った。

4.4 文生成

Ac 節点を頂点とするある部分グラフが与えられたとき、4.3 節で述べたテンプレート選択によって、入力に対する単一のテンプレートが得られる。その後、以下の手順で本システムの出力となるレシピ文が得られる。

- (1) 4.3 節 (1) で得た r-NE 列群の中から、テンプレート内の r-NE タグの順に並ぶものを選択する。
- (2) テンプレート内の r-NE タグを、入力により与えられた r-NE 列に対応する単語列で置き換える。

(1) において、部分グラフの木構造の入替えが可能な場合、その各々のキーの順序でテンプレートに代入しコストを評価している。たとえば、“切/Ac (じゃがいも/F 人参/F)” といった木構造の場合、それに対するキー群は、“F-F-Ac” (じゃがいも-人参-切) と、“F-F-Ac” (人参-じゃがいも-切) が得られるため、“F-F-Ac” をキーとするテンプレートに両方の単語列を代入し、コストを計算している。

5. テンプレート選択基準

学習テキストから自動生成されたテンプレートは膨大であり、入力されたフローグラフに複数のテンプレートがマッチすることが多い。また、自動解析されたテンプレートの中には認識されなかった内容語を含むなど、文生成に適さないテンプレートが存在する。そこで本研究では、最適なテンプレートを選択するためのコストを定義し利用する。コストの計算には、負対数確率 (5.1 節)、内容語残存コスト (5.2 節)、エントロピー (5.3 節) を用いる。最後に、テンプレート選択コスト (5.4 節) でこれら 3 つのコストを統合し、1 つのテンプレートに対するコストを計算する。

5.1 負対数確率

頻出するテンプレートほど妥当なテンプレートと考えられるので、これをコストに反映する。各テンプレート t_i の負対数確率によるコスト $L(t_i)$ を、学習されたテンプレート全体に対するテンプレート t_i の出現確率で計算する。

$$L(t_i) = -\log_2 \left(\frac{F(t_i)}{\sum F(t_i)} \right) \quad (1)$$

ここで $F(t_i)$ は、学習コーパス内で t_i の出現回数を示し、 $L(t_i)$ は小さいほうが頻出するテンプレートであり、より文生成に適していることが期待される。

5.2 内容語残存コスト

4.2 節で作成したテンプレートプールの中には、次のよ

^{*2} <http://plata.ar.media.kyoto-u.ac.jp/kadowaki/>

うな内容語を含むテンプレートが存在する.

- (1) r-NE 認識器の誤り
- (2) r-NE ではない内容語

(1) の例としては,

“最後に /塩/F で /味/Sf を /調え/Ac する。”
 があげられる. この例では, “最後” という内容語が r-NE として認識されずに残存している. こういった内容語が残存したテンプレートを選択することを抑制するために, 残存している内容語の個数に応じてテンプレートに対してコストを与える. (2) は, たとえば調理者の動作を修飾する副詞があげられる. たとえば, “好み”, “よく” が残存したテンプレートの例を示す.

- 好みで F を Ac する。
- よく F を Ac する。

内容語が残存しているかどうかは, 形態素解析を行った際の品詞で判別する. 解析器 (KyTea [19]) が採用している品詞体系 [20] に従って, 以下のとおり内容語, およびテンプレートの骨組みに必要な機能語を定義した.

内容語: 名詞, 動詞, 副詞, 形容詞, 形状詞, 接尾辞, 接頭辞, 連体詞, 代名詞, 接続詞, 感動詞, URL, 英単語, 言いよどみ, ローマ字文

機能語: 助詞, 補助記号, 語尾, 助動詞, 空白, 記号

テンプレート作成過程における自動品詞付与後 (図 4 の上から 2 番目) と重要語タグで置換後 (図 4 の上から 4 番目) の内容語, 機能語, r-NE の割合を表したものが表 2 である. r-NE タグで置き換えたテンプレートの中に, 依然として 10.3% 内容語が残存しており, 内容語残存について考慮する必要があることが分かる.

テンプレート t_i に対する内容語が残存しているコスト $C(t_i)$ を以下の式で定義する.

$$C(t_i) = N \quad (2)$$

ここで, N は t_i の中に含まれる内容語の個数を表す. 内容語残存コスト $C(t_i)$ が小さいほど, 内容語を含まないテンプレートを選択していることが期待される.

5.3 エントロピー

単純にテンプレート頻度に従ってテンプレートを選択すると, 以下の例のように, 活用語尾や助詞が不適切なテン

表 2 自動品詞付与後と重要語タグで置換後の内容語, 機能語, r-NE の割合

Table 2 The ratio of content words, function words and r-NE after giving word automatically and after replacing each word with keyword tag automatically.

	品詞付与後	重要語タグ置換後
内容語	44.6%	10.3%
機能語	55.4%	55.2%
r-NE	-	34.5%

プレートが選択されることがある.

卵/F を ほぐ/Ac する。
 イカ/F に 皮/F を む/Ac する。

この問題に対処するために, 単語 n-gram モデルによるエントロピー [21] を導入する. エントロピーの導入により, 生成文内の自然な単語接続を評価することができる. また, 活用語の活用形や助詞を選択するための特別なメカニズムが不要となる.

以下の手順で, エントロピーによるコストを求める.

- (1) 収集した手順書から単語 3-gram モデル [22], [23] を構築する.
- (2) 各テンプレート t_i に r-NE の単語列を代入し, 文 $S_i = \text{BT} \cdot \text{BT} \cdot w_1 w_2 \cdots w_n \cdot \text{BT}$ を生成する. ここで, BT は文区切りを表す特殊記号である.
- (3) 以下の式で文 S_i における文字あたりのエントロピー $H(t_i)$ を計算する.

$$H(t_i) = \frac{1}{\sum_{k=1}^{n+1} |w_k|} \sum_{k=1}^{n+1} -\log_2 P(w_k | w_{k-2} w_{k-1}) \quad (3)$$

ここで, $|w_k|$ は, 単語 w_k の文字数を表す.

$w_1 w_2 \cdots w_n$ は $H(t_i)$ が小さいほど学習コーパス中に現れやすい単語の並びとなるため, より自然な文であることが期待される.

5.4 テンプレート選択コスト

5.1-5.3 節によって得られたコストを用いて, 各テンプレートから生成された文 t_i のコスト $Cost(t_i)$ を以下の式で計算する.

$$Cost(t_i) = W_L \cdot L(t_i) + W_C \cdot C(t_i) + W_H \cdot H(t_i) \quad (4)$$

ここで, W_L, W_C, W_H は, $L(t_i), C(t_i), H(t_i)$ に対する重みであり, 学習コーパスを用いてグリッドサーチで決定する. 各 t_i に対して, 式 (4) により $Cost(t_i)$ を計算し, 式 (5) により, 最も選択コストが低いテンプレート \hat{t}_i を選択する.

$$\hat{t}_i = \underset{t_i}{\operatorname{argmin}} Cost(t_i) \quad (5)$$

6. 評価実験

本システムはフローグラフを入力としてレシピ文章を生成する. 本章では, 提案手法の有効性を確かめるために, 手順文書の生成を行い, 生成された文を評価する.

6.1 実験環境

テンプレート作成 (4.2 節), およびエントロピー計算 (5.3 節) のための言語モデル構築に用いる学習コーパス

表 3 コーパスの諸元

Table 3 Corpus specifications.

用途	レシピ数	文数	r-NE 数	単語数	文字数
F_{test}	40	162	1,337	3,709	18,254
F_{dev}	160	674	6,539	18,742	94,249
NER 学習	360	2,813	12,101	51,847	97,911
テンプレート学習	429,483	2,173,516	17,613,471	51,099,450	254,124,935

表 4 各学習コーパスサイズに対するグリッドサーチの結果 (BLEU 上位 10 通りの組合せ)

Table 4 A result of grid search for each corpus size (A list of top 10 BLEU values).

学習コーパスの文数															
33,961 (1/64)				135,844 (1/16)				543,379 (1/4)				2,173,516 (1/1)			
BLEU	W_L	W_H	W_C	BLEU	W_L	W_H	W_C	BLEU	W_L	W_H	W_C	BLEU	W_L	W_H	W_C
14.25	0.2	0.9	0.6	17.85	0.2	0.8	0.5	21.80	0.2	0.9	0.4	23.65	0.1	0.5	0.2
14.23	0.2	0.8	0.5	17.81	0.2	0.7	0.4	21.79	0.2	1.0	0.5	23.61	0.1	0.6	0.3
14.23	0.1	0.6	0.4	17.78	0.2	0.9	0.5	21.73	0.1	0.5	0.3	23.59	0.1	0.6	0.2
14.22	0.3	1.0	0.7	17.76	0.1	0.4	0.2	21.73	0.2	1.0	0.6	23.58	0.1	0.6	0.4
14.22	0.1	0.5	0.3	17.76	0.2	0.8	0.4	21.70	0.2	0.9	0.3	23.46	0.1	0.5	0.1
14.22	0.2	1.0	0.6	17.75	0.2	0.7	0.5	21.69	0.2	0.9	0.5	23.43	0.1	0.5	0.3
14.20	0.1	1.0	0.9	17.70	0.1	0.8	0.5	21.69	0.1	0.5	0.2	23.37	0.1	0.6	0.1
14.20	0.1	1.0	0.7	14.20	0.1	0.7	0.5	21.69	0.2	1.0	0.4	23.37	0.1	0.6	0.5
14.20	0.2	1.0	0.7	17.68	0.1	0.6	0.4	21.66	0.2	0.8	0.3	23.28	0.1	0.7	0.4
14.20	0.1	0.9	0.7	17.65	0.1	0.4	0.3	21.66	0.1	0.6	0.3	23.26	0.1	0.7	0.5

として 429,483 レシピ (2,173,516 文) を収集した (表 3)。また、入力とするフローグラフの元となった 200 レシピがフローグラフに変換されたコーパスが存在し [2]、このフローグラフコーパスのレシピは学習コーパスからは除かれている。このうち 160 レシピ (F_{dev}) を用いて、コスト計算 (5.4 節) で各重み W_L , W_H , W_C を決定した。残りの 40 レシピ (F_{test}) をテストコーパスとして入力に与えて文を生成した。生成した文の評価は BLEU [15] ($N = 4$) による自動評価と主観評価の 2 種類を行った。BLEU では、出力されたレシピ文章とフローグラフに付随している人が記述したレシピ文章 (参照文章) とを形態素 n-gram でなお、機械翻訳で用いる BLEU では、1 文を評価単位とするが、本研究では 1 レシピを評価単位とした。生成文の例は付録に示す。

6.2 グリッドサーチによるテンプレート選択コストの重み探索

テンプレートのコスト計算 (5.4 節) で用いる重み W_L , W_H , W_C を、 F_{dev} を開発セットとしてグリッドサーチで求めた。グリッドサーチの探索範囲は、 $W_L, W_H, W_C \in \{0.1, 0.2, 0.3, \dots, 1.0\}$ とした。この最適な重み探索の目的関数としては、開発セットにおける全部分グラフに対する BLEU 値の平均を用いた。また、学習コーパスサイズによる影響を考慮するために、学習コーパスを 1/1, 1/4, 1/16, 1/64 と変化させて、それぞれのコーパスサイズでパラメー

タの重み探索を行った。なお、4.2 節で述べたとおり、テンプレート選択では、(1) で得たキー群に対して、最初にテンプレートを発見できたキーのみ (2) の処理を行うようにした。

表 4 に学習コーパスサイズごとに目的関数とした BLEU が高い上位 10 件の重みの組合せを表す。いずれのコーパスサイズにおいても、上位の組合せは $W_L : W_H \approx 1 : 4 \sim 1 : 5$, $W_L : W_C \approx 1 : 2 \sim 1 : 3$ となっており、重み推定に用いるコーパスのサイズに依存しないことが確認できた。

6.3 テンプレートカバレッジ

どの程度の学習コーパスでどの程度のテンプレート事例がカバーされているのかを確認するため、 F_{test} および F_{dev} (フローグラフを構築した 200 レシピ, 2,938 部分グラフ) の各部分グラフに対して、4.3 節 (2) で得られたテンプレート群のカバレッジを学習コーパスの量に応じて計算した。

表 5 に学習コーパスの量に応じたカバレッジテンプレート群の中に含まれるテンプレートの個数の平均 (平均適合テンプレート数)、入力の部分グラフに対してテンプレート群が得られなかった回数 (不適合部分グラフ数)、入力の部分グラフに対して、マッチするテンプレートが 1 つ以上あった割合 (カバレッジ) を示す。この結果から、今回利用した 429,483 レシピの学習コーパスによって、99.6% の部分グラフに対しテンプレートをあてはめることができ

表 5 各学習コーパスサイズに対する統計結果

Table 5 Statistical results of various template pool sizes.

テンプレート学習コーパスの文数	33,961 (1/64)	135,844 (1/16)	543,379 (1/4)	2,173,516 (1/1)
総テンプレート数	31,602	124,582	487,685	1,894,229
平均適合テンプレート数	784	1,570	5,071	18,697
不適合部分グラフ数	88	51	29	13
カバレッジ [%]	97.0	98.3	99.0	99.6

表 6 10 レシピに対する主観評価

Table 6 Subjective assessment for 10 recipes.

BLEU	評価者 1	評価者 2	評価者 3	評価者 4	評価者 5	評価者 6	評価者 7	平均
9.31	3	4	3	2	5	3	4	3.43
13.26	2	4	2	5	5	3	4	3.57
18.14	3	4	3	4	4	3	4	3.57
23.07	3	3	3	1	5	2	4	3.00
27.88	3	5	3	3	4	4	4	3.71
32.48	4	5	5	5	5	3	5	4.57
37.24	5	4	5	4	5	4	5	4.57
38.33	5	5	5	5	5	4	5	4.86
43.57	4	3	3	4	5	4	5	4.00
50.90	5	5	5	5	5	4	5	4.86
平均	3.70	4.20	3.70	3.80	4.80	3.40	4.50	4.01
PCC	+0.85	+0.25	+0.70	+0.46	+0.25	+0.67	+0.86	+0.78

PCC はピアソン相関係数を表す。

ることが分かった。また、この 1/4 の学習コーパスでも 99.0% の部分グラフをカバーすることができ、ある程度の学習コーパスがあれば、活用語尾をテンプレートに含めた場合でも十分に必要な表現がカバーされることが分かる。

6.4 被験者実験による主観評価

本研究では自動評価指標として BLEU を用いているが、BLEU による評価値は生成された文章と参照文章との形態素 n-gram の一致度を評価したものであり、生成された文の意味的・文法的な正しさを十分に反映していない可能性がある。たとえば、調理手順の場合、動作に対する材料の順序は可換であることがよくあるが、参照文章とは異なる順序で文生成すると BLEU は低下する。そのため、被験者実験によって生成された文の評価を行うとともに、BLEU との相関を測って BLEU による評価の妥当性を検証する。本論文では、生成された文章から、参照文章と同等の手続きを読者が読み取れるかどうかを、生成された文章の品質を定める基準とする。7 名の評価者（20–23 歳の男性 5 名、女性 2 名。全員が日本語を母語とし、週 1 回以上料理を行い、情報学以外を専門とする学生）に様々な BLEU の 10 個のフローグラフから出力された文章と参照文章を提示し、以下の質問に回答させた。

- Q. 生成された文に従って調理を行った場合、参照文章とどれくらい同じ料理が作れますか？
- A. 5：完全に、4：ほとんど、3：部分的に、2：異なる、

1：調理不可能

なお、実験条件は、後述の 6.5 節における、尤度選択方法 (F+C+E) で、学習コーパス量は 1/1 を用いて文生成を行った。生成されたレシピには平均 11.3 文含まれていた。表 6 に主観評価の結果を示す。評価値の平均は 4.01 であり、平均的にはおおむね適切な文生成が行われていることが確認できる。また、いずれの評価者においても、評価値が BLEU と強い相関があり、BLEU が文生成の質を測る有効な評価指標であることが確認できる。また、BLEU が 32.48 以上であるもののうち、91.4% が「5：完全に」、または「4：ほとんど」、と評価されていた。このことから、おおそ BLEU が 33 以上の場合、適切な文が生成され、人が書いたものと同等の文章を再現できているといえる。

本研究では料理経験がある人を対象としているが、料理経験がほとんどない人を対象としても同等の結果が得られると推察できる。なぜなら、生成文の重要語は、参照文の重要語と同一であり、仮にその重要語の意味・存在を知らなくても、参照文と同様の述語項構造が選択されているか否かで評価を行うことができるからである。

6.5 BLEU による自動評価

学習コーパスサイズごとのグリッドサーチ (6.2 節) で得た最適な重みを用いて、 F_{test} を入力として与えて得られた文を BLEU で評価した。テンプレート選択 (4.3 節) で用いた 3 つのコスト (負対数確率、内容語残存コスト、エ

表 7 コーパスサイズに対する各手法の BLEU

Table 7 BLEU of each method for each corpus size.

学習コーパス の文数	33,961 (1/64)	135,844 (1/16)	543,379 (1/4)	2,173,516 (1/1)
L	11.16	15.47	18.80	18.98
C	11.32	13.24	17.47	18.13
H	12.95	15.31	20.57	23.89
L+C	11.19	13.23	18.20	18.63
L+H	14.06	20.94	27.23	32.62
C+H	19.67	23.72	27.42	28.67
L+C+H	19.35	23.91	29.44	33.35

ントロピー) の文生成精度への寄与を測るために、以下の
ように 7 通りの尤度選択手法により文生成を行った。

L: 負対数確率のみ

C: 内容語残存コストのみ

H: エントロピーのみ

L+C: 負対数確率+内容語残存コスト

L+H: 負対数確率+エントロピー

C+H: 内容語残存コスト+エントロピー

L+C+H: 負対数確率+内容語残存コスト
+エントロピー

それぞれの手法において、学習コーパスを 1/1, 1/4,
1/16, 1/64 と変化させて、その学習コーパスサイズに対応
した重みを使用して文生成を行った。評価の際は学習時と
異なり、キーにマッチするすべてのテンプレートを選択対
象とする。

表 7 に BLEU による自動評価の結果を示す。この結果
から、適切なテンプレート選択にはエントロピーと負対数
確率による寄与が大きく、内容語残存コストによる寄与は
あまり大きくないことが分かった。理由としては、エント
ロピーの導入によって活用語尾や適切な助詞の接続がなさ
れるようになったことが、BLEU が評価する n-gram の一
致に合致したこと、負対数確率の導入によって頻出するテ
ンプレートにはより適切ないい回しが含まれていることが
期待できることがあげられる。加えて、内容語残存コスト
の導入は意味内容を変える致命的な文生成を時折避けるこ
とにとどまり、BLEU に対する影響は低いことが理由と
して考えられる。エントロピーを導入しない手法 (L, C,
L+C) は、学習コーパスを増やしても評価値の向上幅が
少なく、学習データに含まれるノイズの影響を受けている
と考えられる。それに対してエントロピーを導入した手法
(H, L+H, C+H, L+C+H) は、コーパスサイズに応じて
評価値が向上しており、適切なテンプレートの選択に寄与
していると考えられる。

BLEU での評価は、「じゃがいもを切ります」と「じゃ
がいもを切る」といった、レシピの作業内容は異ならな
いが、表現が異なる場合の差を考慮していない。しかし、
6.4 節により BLEU による評価結果と人手による評価結果

表 8 活用語の語幹と活用語尾間の接続正解率

Table 8 Junctional accuracy between stem and ending in a
conjugational word.

レシピ数	Ac 数	接続正解数	接続正解率
40	518	493	0.951

に相関を確認できたことが重要である。これにより、今後
のレシピ文生成の研究における評価で、コストのかかる人
手を用いずに、自動評価で生成文の評価を行うことができ
るといえる。

6.6 活用語の活用語尾の正解度

本研究で提案する活用語尾をテンプレート側を含めると
いう設計が機能していることを確認するために、活用語の
語幹と活用語尾の接続の文法性を検証した。 F_{test} を入力と
して、得られた 40 レシピの文中に含まれる、各々の Ac 部
で置き換えられた単語列と、テンプレート側の活用語尾と
の接続の文法的正しさを人手によって判断した。なお、文
選択は、6.5 節における、尤度選択方法 (F+C+E) で、学
習コーパス量は 1/1 の場合の生成結果を用いた。表 8 に
その結果を示す。実験結果から、95.1%の精度で活用語と
活用語尾との接続が文法的に正しいことが確認できた。こ
の結果から、活用語尾をテンプレートに含めるという枠組
みが、適切に動作することが確認された。

7. 関連研究

言語生成の研究は、近年では小説の生成 [24] や詩の生
成 [25], [26] などが行われているが、これらのドメインでは
人物の感情や作者の意図といった表面上の文面には現れな
い情報が多く、同じ内容を述べる場合でも表現方法が異なる。
これに対して本研究では、手順書を対象としている。手
順書は小説や詩などと比べ客観的な文書であると考えら
れ、そのため、何かを述べるときにその書かれ方の幅が比
較的狭いと考えることができる。したがってということが決
まっている中で、その肉付けを行う本研究の対象として適
切であると考えられる。

本研究では、意味表現としてのフローグラフが骨子とし
て与えられた中で、それにどのような機能語、接続詞、語
尾などを肉付けするかということに注目した。こうした枠
組みは音声対話システムなどの応答生成にも必要である
が、これまではあらかじめ定義したテンプレートを利用し
た文 [27], [28], [29] やあらかじめ定義・収集した文の中
で最も一致度の高い文 [30], [31] を読み上げていた。しかし、
あらかじめ定義したテンプレートを利用した手法はテン
プレートの作成にコストがかかることや、一致度の高い文
の読み上げを行う手法は表現方法に自由度がないという問
題があった。これに対して本研究では、コンピュータが内部
に持つ意味表現をベースとして、大量に収集したコーパス

からテンプレートを作成し統計的な生成を行うことで、自動的なテンプレート作成、および入力に応じた自動的なテンプレート選択によりこれらの問題に対処した。

2章で述べたとおり、手順書からの文生成において重要なのはということといい方である。本研究では、いうことは主に内容語で表現されていると考えて、テンプレート [32] を自動生成しておき、適切に選択する設計とした。

学習データからテンプレートを学習する枠組みは、主に対話システムのための文生成で研究されてきたが [14]、テンプレート作成時に不要な単語や、文脈を生成する接続詞などを含むテンプレートをあらかじめ除去を行い、学習データの文数は 18,466 文、不要なテンプレートを除去後に残ったテンプレート数は 512 文と小規模な学習で行われていた。これに対し、本研究ではテンプレート作成時にフィルタリングを行わず、テンプレート選択時に内容語を含むテンプレートに対してコストを設け、学習データの文数は 2,173,516 文、テンプレート数は 1,894,229 文と、大規模な学習を行った。また、文献 [14] ではユーザの評価値を入力として文を生成するのに対し、本研究ではフローグラフの部分グラフを中間表現として入力し、文書を生成する点も異なる。

文生成は、形式言語理論でいうところの文法を仮定し、それによるアルファベット列を導出する課題とみることができる。この観点では、テンプレート方式は、開始記号からの生成規則を除けば、非終端記号が右辺にない生成規則のみからなる文法を仮定しているといえる。これに対して、確率的正規言語を仮定して将棋のコメントを生成する研究 [33] がある。また、文脈自由文法などを仮定することも可能である [34]。これらの再帰を許す文法では生成される文表現の幅は広いが、同じ単語列が繰り返されるなどの非文を生成する可能性があるという問題がある。本研究では、Web 上の実例をテンプレートとして活用することにより、表現の幅の問題を対処しながら、非文を生成しにくいというテンプレート方式の長所を活かす設計を提案している。さらに、どの程度の実例があれば手順書を生成するのに十分かを実験的に明らかにしている。

本研究の応用は 1 章で述べたように複数考えられるが、なかでも重要な応用例として Smart Kitchen [11] があげられる。Smart Kitchen とは、映像認識技術などを用いた、コンピュータによる調理補助システムであり、調理者に次の調理手順を示したり、動作の説明を行ったりするシステムである。このシステムに対する応用として、指示モジュールからの指示として部分グラフが与えられたとき、その部分グラフから文生成を行い、調理者に示すことができる。

また別の応用例として、インターネットユーザの文章生成の補助システム、フローグラフとして表現された 2 言語間のレシピ翻訳システム、調理映像から自動レシピ文生成システム [9], [10], [35] があげられる。

8. おわりに

本研究では手順を示したフローグラフから手順を示す文を自動生成する手法について述べた。動作対象（主に r-NE タグの食材 F）、あるいは手段や方向（主に r-NE タグの道具 T）と動作動詞を接続する格助詞は、生成された文の意味に大きく影響する。そこで本研究では、内容語にあわせたテンプレートを多くの統計的文生成の手法と同様に、言語モデルを用いて選択することとしている。この手法を活用語の活用形自動学習にも応用しているのが本研究の特徴である。具体的には、レシピドメインにおいて学習コーパスからテンプレートを作成し、入力のフローグラフからコスト計算によってテンプレートを選択した。コスト計算の手法として、負対数確率、内容語残存コスト、エントロピーを用いることを提案した。実験では 429,483 レシピから学習を行い、40 レシピのフローグラフを入力として与え、生成された文とレシピの原文との比較を BLEU で評価し、提案した選択基準が有効であることを確認した。また、10 レシピに対して主観評価を行い、BLEU とレシピの意味的正確さに相関があることを確認した。また、出力文の BLEU が 33 程度を超えると、出力文に従って実際に調理を行うことを想定した場合、元のレシピをほとんど再現できることが示された。

謝辞 本研究は、JSPS 科研費 26280084 および 24240030 の助成を受けたものです。ここに謝意を表します。

参考文献

- [1] 浜田玲子, 井手一郎, 坂井修一, 田中英彦: 料理テキスト教材における調理手順の構造化, 電子情報通信学会論文誌, Vol.J85-DII, No.1, pp.79–89 (2002).
- [2] Mori, S., Maeta, H., Yamakata, Y. and Sasada, T.: Flow Graph Corpus from Recipe Texts, *Proc. 9th International Conference on Language Resources and Evaluation*, pp.2370–2377 (2014).
- [3] 吉野幸一郎, 森 信介, 河原達也: 述語項構造を介した文の選択に基づく音声対話用言語モデルの構築, 人工知能学会論文誌, Vol.29, No.1, pp.53–59 (2014).
- [4] Faria, F., dos Santos, J., Rocha, A. and Torres, R.: Automatic Classifier Fusion for Produce Recognition, *2012 25th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, pp.252–259 (2012).
- [5] Rohrbach, M., Amin, S., Andriluka, M. and Schiele, B.: A database for fine grained activity detection of cooking activities, *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp.1194–1201 (2012).
- [6] 井上 仁, 橋本敦史, 中村和晃, 船富卓哉, 山肩洋子, 上田真由美, 美濃導彦: 食材認識のための画像と食材切断時の振動音及び荷重の利用, 電子情報通信学会論文誌 D, Vol.97, No.9, pp.1490–1502 (2014).
- [7] Naim, I., Song, Y.C., Liu, Q., Kautz, H., Luo, J. and Gildea, D.: Unsupervised Alignment of Natural Language Instructions with Video Segments, *Proc. 28th National Conference on Artificial Intelligence* (2014).
- [8] Rohrbach, M., Qiu, W., Titov, I., Thater, S., Pinkal, M. and Schiele, B.: Translating Video Content to Natural

- Language Descriptions, *Proc. 14th International Conference on Computer Vision* (2013).
- [9] Yu, H. and Siskind, J.M.: Grounded Language Learning from Video Described with Sentences, *Proc. 51st Annual Meeting of the Association for Computational Linguistics* (2013).
- [10] Yamakata, Y., Imahori, S., Sugiyama, Y., Mori, S. and Tanaka, K.: Feature Extraction and Summarization of Recipes using Flow Graph, *Proc. 5th International Conference on Social Informatics*, LNCS, Vol.8238, pp.241–254 (2013).
- [11] Hashimoto, A., Mori, N., Funatomi, T., Yamakata, Y., Kakusho, K. and Minoh, M.: Smart Kitchen: A User Centric Cooking Support System, *Proc. 12th Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pp.848–854 (2008).
- [12] Angeli, G., Liang, P. and Klein, D.: A simple domain-independent probabilistic approach to generation, *Proc. 2010 Conference on Empirical Methods in Natural Language Processing*, pp.502–512 (2010).
- [13] Kondadadi, R., Howald, B. and Schilder, F.: A Statistical NLG Framework for Aggregated Planning and Realization, *Proc. 51st Annual Meeting of the Association for Computational Linguistics*, pp.1406–1415 (2013).
- [14] Higashinaka, R., Prasad, R. and Walker, M.A.: Learning to Generate Naturalistic Utterances Using Reviews in Spoken Dialogue Systems, pp.265–272 (2006).
- [15] Papineni, K., Roukos, S., Ward, T. and Zhu, W.-J.: BLEU: A Method for Automatic Evaluation of Machine Translation, *Proc. 40th Annual Meeting of the Association for Computational Linguistics*, pp.311–318 (2002).
- [16] Dreyer, M. and Eisner, J.: Discovering Morphological Paradigms from Plain Text Using a Dirichlet Process Mixture Model, *Proc. 2011 Conference on Empirical Methods in Natural Language Processing* (2011).
- [17] Neubig, G., Nakata, Y. and Mori, S.: Pointwise Prediction for Robust, Adaptable Japanese Morphological Analysis, *Proc. 49th Annual Meeting of the Association for Computational Linguistics*, pp.529–533 (2011).
- [18] Mori, S., Sasada, T., Yamakata, Y. and Yoshino, K.: A Machine Learning Approach to Recipe Text Processing, *Proc. 1st Cooking with Computer Workshop*, pp.29–34 (2012).
- [19] 森 信介, Graham, N., 坪井祐太: 点予測による単語分割, 情報処理学会論文誌, Vol.52, No.10, pp.2944–2952 (2011).
- [20] 小椋秀樹, 小磯花絵, 富士池優美, 原 裕: 『現代日本語書き言葉均衡コーパス』形態論情報規程集, 独立行政法人国立国語研究所 (2008).
- [21] Ratnaparkhi, A.: 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).
- [22] Brown, P.F., Pietra, S.A.D. and Mercer, R.L.: An Estimate of an Upper Bound for the Entropy of English, *Computational Linguistics*, Vol.18, No.1, pp.31–40 (1992).
- [23] 森 信介, 山地 治: 日本語の情報量の上限の推定, 情報処理学会論文誌, Vol.38, No.11, pp.2191–2199 (1997).
- [24] 高木大生, 佐藤理史, 駒谷和範: 会話を中心とした超短編小説の自動生成, *The 28th Annual Conference of the Japanese Society for Artificial Intelligence*, 2014 (2014).
- [25] Agirrezabal, M., Arrieta, B., Astigarraga, A. and Hulden, M.: POS-Tag Based Poetry Generation with WordNet, *Proc. 14th European Workshop on Natural Language Generation*, pp.162–166 (2013).
- [26] Zhang, X. and Lapata, M.: Chinese poetry generation with recurrent neural networks, *Proc. 2014 Conference on Empirical Methods in Natural Language Processing*, pp.670–680 (2014).
- [27] Langkilde, I. and Knight, K.: Generation That Exploits Corpus-based Statistical Knowledge, *ACL98*, pp.704–710 (1998).
- [28] Walker, M.A., Rambow, O.C. and Rogati, M.: Training a sentence planner for spoken dialogue using boosting, *Computer Speech & Language*, Vol.16, No.34, pp.409–433 (2002).
- [29] Komatani, K., Ueno, S., Kawahara, T. and Okuno, H.G.: Flexible Guidance Generation Using User Model in Spoken Dialogue Systems, *Proc. 41st Annual Meeting of the Association for Computational Linguistics*, pp.256–263 (2003).
- [30] Yoshino, K., Mori, S. and Kawahara, T.: Spoken dialogue system based on information extraction using similarity of predicate argument structures, *ACL SIGDIAL*, pp.59–66, Association for Computational Linguistics (2011).
- [31] 稲葉通将, 神園彩香, 高橋健一: Twitter を用いた非タスク指向型対話システムのための発話候補文獲得, 人工知能学会論文誌, Vol.29, No.1, pp.21–31 (2014).
- [32] Reiter, E.: NLG vs. Templates, *Proc. 5th European Workshop on Natural Language Generation*, pp.147–151 (1995).
- [33] 亀甲博貴, 三輪 誠, 鶴岡慶雅, 森 信介, 近山 隆: 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).
- [34] Konstas, I. and Lapata, M.: Concept-to-text Generation via Discriminative Reranking, *Proc. 50th Annual Meeting of the Association for Computational Linguistics*, pp.369–378 (2012).
- [35] Regneri, M., Rohrbach, M., Wetzell, D., Thater, S., Schiele, B. and Pinkal, M.: Grounding Action Descriptions in Videos, *Trans. Association for Computational Linguistics*, Vol.1, pp.25–36 (2013).

付 録

生成文例 6.4 節で評価させた生成文章のうち, BLEU が 9.31, 33.52, 50.90 であるものを示す。

BLEU : 9.31 の生成文

アク/F をと/A c る。
 たっぷり/Q の水/F と モツ/F をゆで/Ac る。
 水/F を替え/Ac る。
 繰り返/Ac します。
 ごぼう/F は ささがき/Sf にし/Ac ます。
 浸/Ac します。
 鍋/T に水/F と モツ/F を入れ/Ac る。
 にこみ/Ac ます。
 ごぼう/F に 火/Sf が 通/Af ったら、味噌/F を入れ/Ac る。
 煮込/Ac みます。

BLEU : 33.52 の生成文

ご飯/F は洗/Ac っておく。
ザル/T に上げ/Ac する。
焼き鮭/F は骨/F を取り除/Ac きます。
ほぐ/Ac します。
フライパン/T にスープ/F とご飯/F を入れ/Ac する。
中火/T にかける/Ac する。
鮭/F が沸騰/Af したら、かぶ/F の葉/F を入れ/Ac します。
たまご/F を溶/Ac きます。
一煮立ち/Af したら、軽く回しかける/Ac する。
パセリ/F を散ら/Ac する。

BLEU : 50.90 の生成文

じゃがいも/F は 1 c m 厚さ/Sf の輪切り/Sf にし/Ac ます。
ゆで/Ac する。
玉ねぎ/F を食べやすい大きさ/Sf に切/Ac する。
牛肉/F と玉ねぎ/F を炒めあわせ/Ac する。
ポテト/F を加え/Ac する。
醤油/F とみりん/F で味付け/Ac する。



山崎 健史 (学生会員)

2014 年京都大学工学部情報学科卒業。
同年同大学大学院情報学研究科修士課程に進学。



吉野 幸一郎 (正会員)

2009 年慶應義塾大学環境情報学部卒業。
2011 年京都大学大学院情報学研究科修士課程修了。
2014 年同博士後期課程修了。
同年日本学術振興会特別研究員 (PD)。
2015 年より奈良先端科学技術大学院大学情報科学研究科特任助教。
京都大学博士 (情報学)。
音声言語処理および自然言語処理、特に音声対話システムに関する研究に従事。
2013 年度人工知能学会研究会優秀賞受賞。IEEE, ACL, 言語処理学会各会員。



前田 浩邦

2012 年京都大学理学部卒業。
2014 年同大学大学院情報学研究科知能情報学専攻修士課程修了。
同年サイボウズ株式会社入社、現在に至る。



笹田 鉄郎 (正会員)

2007 年京都大学工学部電気電子工学科卒業。
2009 年同大学大学院情報学研究科修士課程修了。
同年同大学院博士後期課程に進学。
2012 年同大学院情報学研究科博士後期課程単位取得認定退学。
京都大学学術情報メディアセンター教務補佐員、現在に至る。
自然言語処理に関する研究に従事。言語処理学会会員。



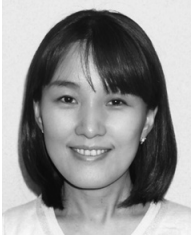
橋本 敦史 (正会員)

2005 年京都大学工学部情報学科卒業。
2006 年経産省 Vulcanus in Europe プログラム国費奨学生。
2011 年京都大学大学院情報学研究科博士後期課程研究指導認定退学。
現在京都大学大学院法学研究科助手。
主にマルチセンシング環境における人間行動観測を介した物体認識、動作認識、物体追跡に関する研究に従事。
京都大学博士 (情報学)。
電子情報処理学会会員。



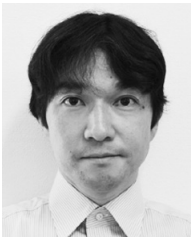
船富 卓哉 (正会員)

2002 年京都大学工学部情報学科卒業。
2004 年同大学大学院情報学研究科修士課程修了。
2006 年日本学術振興会特別研究員 DC2。
2007 年京都大学大学院情報学研究科博士後期課程修了。
2007 年京都大学学術情報メディアセンター助教。
2014 年 Stanford University, Computer Science Department 客員助教。
2015 年奈良先端科学技術大学院大学情報科学研究科准教授、現在に至る。
コンピュータビジョン、マルチメディア、3次元モデル、Computational photography に関する研究に従事。
博士 (情報学)。
ACM, IEEE, 電子情報通信学会各会員。



山肩 洋子 (正会員)

2005年京都大学大学院情報学研究科博士後期課程研究指導認定退学。同年京都大学情報学研究科研究員。2006年(独)情報通信研究機構専攻研究員。京都大学特定講師, 准教授, 特定准教授を経て, 2015年より日本学術振興会特別研究員。電子情報通信学会「食メディア研究会」副委員長。レシピや調理行動に関するマルチメディア処理の研究に従事。電子情報通信学会, 人工知能学会各会員。博士(情報学)。



森 信介 (正会員)

1998年京都大学大学院工学研究科電子通信工学専攻博士後期課程修了。同年日本アイ・ビー・エム(株)入社。2007年より京都大学学術情報メディアセンター准教授。京都大学博士(工学)。1997年情報処理学会山下記念研究賞受賞。2010年, 2013年情報処理学会論文賞受賞。2010年第58回電気科学技術奨励賞受賞。言語処理学会, データベース学会, ACL各会員。