

フローグラフからのレシピ文自動生成

山崎 健史^{1,a)} 森 信介^{2,b)} 河原 達也^{2,c)}

概要: 調理動作からレシピ文の自動生成を目標として、フローグラフからのレシピ文自動生成に関する研究が進められている。本研究では、レシピ文の言語モデルを用いた手法を提案する。元のレシピに対する BLEU 値で評価した。

1. はじめに

コンピュータが日々の生活に浸透していくにしたがって、自然言語生成技術の需要が高まっている。例えば、コンピュータが手順文書を理解し、それを実施している作業の状況を認識できるとすると、次に何をすべきかを提示することが可能となるであろう。そのような状況下では、次の作業内容を文で画面に提示するかスピーカーから発声することになる。

このような背景の下、手順書の意味表現 (フローグラフ)[1][2] からの文生成が研究されている [3]。手順文書の意味表現は、人が直接描くことも考えられるが、複数の手順書の意味表現に対する演算結果や映像認識の結果として得られることが想定される。意味表現に対する演算としては、同じ料理名の複数のレシピの意味表現から計算される典型的な手順の意味表現 [4] などがある。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 [5] [6] 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 [7] などがある。

本論文では、フローグラフで表現されたレシピ [2] から自然言語のレシピを生成する方法について述べる。先行研究として、テンプレートの自動獲得とそれを用いた文生成 [3] がある。この手法では、活用語と語尾が合わないテンプレートを選択したり、重要語の認識誤りが主な原因でテンプレート中に残存する内容語が意図せず生成文に含まれるなどの問題がある。これらの問題に対処するために、本論文では、言語モデルや内容語残存コストを組み入れることを提案する。実験では、実際にフローグラフから文生

表 1 固有表現の種類

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	

成を行い、提案手法の採否による出力文の質を元のレシピを参照文とした BLEU [8] によって評価する。

2. レシピコーパス

本システムで入力及び学習コーパスとして用いるレシピコーパスは、タイトル、材料一覧、調理手順によって構成される文章群である。レシピコーパスは、レシピサイトからクロールして入手したものである。レシピコーパスの中で、調理手順 (図 2) を解析・編集することで入力及び学習コーパスを得ている。

レシピ文の意味を表現する方法として、フローグラフが存在する (図 1)。フローグラフとは、文章の構造を単語毎の関係で表した無閉路有向グラフである。各節点は、調理を行う際における重要な単語と、レシピ固有表現タグ (以下 r-NE タグ) の組で表現されたレシピ固有表現 (以下 r-NE) が対応している。r-NE タグとは、調理を行う上での各単語の役割を表している (表 1)。枝の矢印は r-NE の関係を表し、また、枝にはその関係に関するラベルがふられている。例えば、図 1 中の上部で、“熱/Ac” 節点は両手 “両手鍋/T” 節点と “T-comp” ラベルによる枝で接続されているが、これは、“熱” という動作は “両手鍋” という道具と道具による補語関係にある、という関係を示している。しかし、本研究ではこの枝のラベルは不必要なため用いない。

¹ 京都大学情報学研究科
京都府京都市左京区吉田本町

² 京都大学学術情報メディアセンター
京都府京都市左京区吉田本町

a) yamasaki@ar.media.kyoto-u.ac.jp

b) forest@i.kyoto-u.ac.jp

c) kawahara@i.kyoto-u.ac.jp

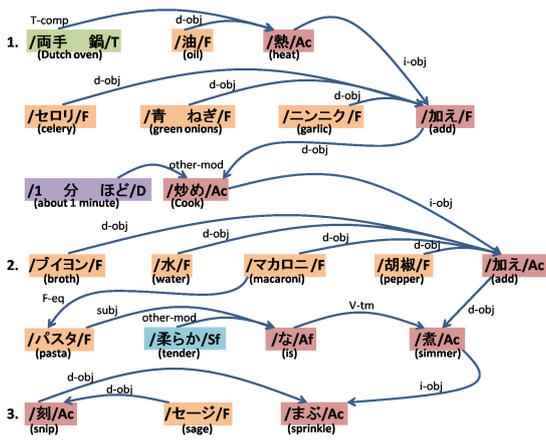


図 1 レシピテキストのフローグラフ

1. 両手鍋で油を熱する。
セロリと青ねぎとニンニクを加える。
1分ほど炒める。
2. ブイヨンと水とマカロニと胡椒を加えパスタが柔らかくなるまで煮る。
3. 刻んだセージをまぶす。

図 2 レシピの調理手順例

熱 Ac((T-COMP 両手鍋 T)(D-OBJ 油 F))
 加え Ac((D-OBJ セロリ F)(D-OBJ 青ねぎ F)
 (D-OBJ ニンニク F))
 炒め Ac((OTHER-MOD 1分ほど D))
 加え Ac((D-OBJ ブイヨン F)(D-OBJ 水 F)
 (D-OBJ マカロニ F)(D-OBJ 胡椒 F))

図 4 入力フォーマットの例

温かいご飯にかけます。
 少し冷めたら、布巾をとる。
 レシピテキスト原文

温かいご飯にかけます。
 少し冷めたら、布巾をとる。

(1)KyTea による単語分割

温か/Sf いご飯/F に かけ/Ac ます。
 少し 冷め/Af たら、 布巾/F を と/Ac る。

(2)PWNER による r-NE 付与

Sf い F に Ac ます。
 少し Af たら、 F を Ac る。

(3)r-NE タグで置き換えてテンプレート生成

図 5 テンプレート生成

3. レシピ文自動生成システム

本章では、本研究で提案するレシピ文自動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を生成するシステムである。本システムは大きく2つの行程に別れており、(1)テンプレート生成(3.2節)、(2)テンプレート選択(3.3節)、レシピ文生成(3.4節)からなる。

3.1 入力

本システムに与える入力は、レシピテキストの意味表現であるフローグラフの部分木である。本研究では、レシピテキストからフローグラフへ変換済みのコーパスを利用する[2]。フローグラフの形式は、図3で表される様に、csv形式で表現される。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最大部分木を構築する。この時、一度使用した節点にはフラグを設定し、2回以上使用しないようにする。

以上の手順によって得られる入力フォーマットの例を図4に示す。

3.2 テンプレート生成

レシピ文生成に用いるテンプレートを学習コーパスから以下の手順で作成する[3]。

- (1) レシピ文をテキスト解析ツール KyTea[9] を用いて単

語単位に分割し、品詞を付与する*1。

- (2) 単語列に対して PWNER[10] を用いて r-NE タグを付与する*2。

- (3) r-NE を r-NE タグで置き換える。

なお、各テンプレートはテンプレートのキーで参照できるようにする。テンプレートのキーとは、テンプレート中に出現する r-NE タグを並べたものである。例えば、図5(3)の第一文のテンプレートのキーは“Sf-F-Ac”となる。以上の一連の流れを図5に示す。

3.3 テンプレート選択

3.2節で作成したテンプレート群の中には、与えられた入力に対して埋められるテンプレートは複数存在し、このままでは一意に特定できない。本節では、尤度をもってテンプレートを選択する手法を紹介する。以下の流れでテンプレートを選択する。

- (1) 入力の係り受け構造から、テンプレートを参照するキーを生成する。この際、図6のように木構造の入れ替えが可能な場合、キー群として得る。
- (2) 各キーに対して、
 - (a) キーの r-NE タグの順に表れるテンプレート群を得る。
 - (b) 各テンプレートに対する尤度(後述)を計算する。
- (3) (2)で計算した尤度が最小となるテンプレートを選択する。

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

*2 <http://plata.ar.media.kyoto-u.ac.jp/kadowaki/index.html>

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	ID	出現箇所	固有表現	単語列	主語(カ格)	対象(ヲ格)	方向(ニ格)	食材	道具	食材	食材	食材	道具	道具	動詞	動作のタイ	その他の備考	
2	ID	char-positi	Ntype	wordseq	subj	d-obj	i-obj	F-by	T-by	F-equal	F-part-of	F-set	T-equal	T-part-of	V-equal	V-timing	othermod	remark
3		1	001-01	-OCT	鍋													
4		2	001-01	-OCF	水													3
5		3	001-01	-OCQ	三カップくらい													
6		4	001-01	-01Ac	い		2		1									
7		5	001-01	-01F	だしの素													
8		6	001-01	-02F	昆布													
9		7	001-01	-02Ac	いれ		5,6		4									

図 3 フローグラフの形式

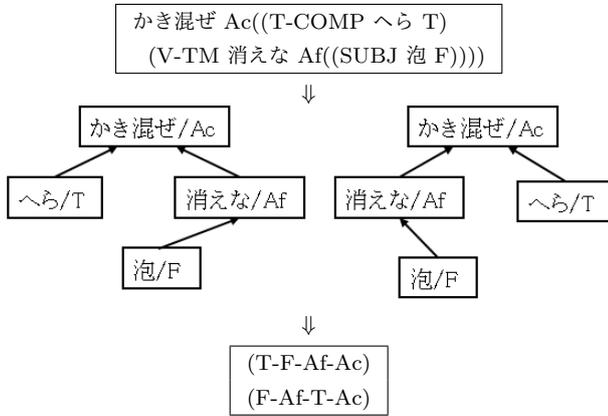


図 6 r-NE 列群の生成

先行研究 [3] ではテンプレートの尤度計算にテンプレート頻度のみを採用していたが、本研究では (1) 頻度 (3.3.1), (2) 内容語残存コスト (3.3.2), (3) エントロピー (3.3.3), を用いてテンプレートの尤度を計算する。

3.3.1 頻度

頻出するテンプレートほど妥当なテンプレートと捉えて、優先度を上げる。各テンプレート T_i の頻度による尤度 L_i を以下の式で計算する。

$$L_i = -\log_2 P\left(\frac{Freq(T_i)}{\sum Freq(T_i)}\right) \quad (1)$$

ここで $Freq(T_i)$ は、学習コーパス内で T_i の出現回数を示す。 L_i は低いほうが良いことが期待される。

3.3.2 内容語残存コスト

3.2 節で作成したテンプレート群の中には、調理手順の内容に関わる単語にもかかわらず、r-NE が割り当てられることなくテンプレートの中に残存してしまったテンプレートが存在する。このようなテンプレートが選択されてしまった場合、意味的に原文とは異なる文を生成しうる。先行研究 [3] のテンプレート頻度のみでは、例えば

”短/St °C で、こんがりする /Ac る。”
 といった” °C ”, ”こんがり” という内容語が含まれたテンプレートを選択するケースが存在していた。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、残存したテンプレートに対してコストを与える。

内容語の当否については、形態素解析を行った際の品詞で判別する。解析器 (KyTea[9]) が規定している品詞体系 [11] に従って、内容語、及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、機能語の品詞を以下に提示する。

内容語： 名詞、動詞、副詞、形容詞、形状詞、接尾辞、接頭

表 2 学習コーパス全体と、r-NE 部分を除いたコーパスにおける内容語と機能語の割合

品詞群 \ 統計コーパス	all	without r-NE
内容語	21.7%	15.7%
機能語	78.3%	84.3%

辞、連体詞、代名詞、接続詞、感動詞、URL、英単語、言いよび、ローマ字文

機能語： 助詞、補助記号、語尾、助動詞、空白、記号

学習コーパス全体と、r-NE を除いたコーパスにおける内容語と機能語の割合を表したものが表 2 である。r-NE で置き換えたテンプレートの中に、依然として 15.7% 内容語が残っており、内容語残存について考慮する必要性が窺える。

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_i を以下の式で定義する。

$$C_i = N \quad (2)$$

ここで、 N は、 T_i 中に存在している内容語の個数である。

3.3.3 エントロピー

先行研究 [3] ではテンプレート頻度のみによる選択をしていたため、

卵/F を ほぐ/Ac る。

といった文末表現が誤った文章を生成してしまう。この問題に対処するために、言語モデルによるエントロピーを導入する。以下の手順で、エントロピーによる尤度を求める。

- (1) 学習コーパスから単語 3-gram モデルを生成する。
- (2) 各テンプレート T_i に r-NE の単語列を代入し、文 $S_i = BTw_1w_2 \dots w_nBTBT$ 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 に対するエントロピー H_i を計算する。

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

H_i が低いほど、 $w_1w_2 \dots w_n$ は学習コーパス中に現れやすい単語の並びであり、より自然な文であることが期待される。

3.3.4 尤度計算

3.3.1 - 3.3.3 項によって得られた尤度を用いて、各テンプレート T_i の尤度 $Cost_i$ を以下の式で計算する。

$$Cost_i = W_L \cdot L_i + W_H \cdot H_i + W_C \cdot C_i \quad (4)$$

ここで、 W_L, W_H, W_C は、 L_i, H_i, C_i に対する重みである。

3.4 文生成

3.3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。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) 3.3節(1)で得たNE列群の中から、テンプレート内のr-NEの順に並ぶものを選択する。
- (2) テンプレート内のr-NEを、選択したr-NE列に対応する単語で置き換える。

4. 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、生成された文の精度を評価する。

4.1 実験環境

テストコーパスとして、40個のフローグラフを提案したシステムの入力として与え、出力されたレシピ文とフローグラフに付随している人が記述したレシピ文(参照文)との比較した。3.2節のテンプレート作成時のコーパス、及び3.3.3項でエントロピー計算のための言語モデルのコーパスは429,691レシピ(1,189,794)文である。比較の評価にはBLEU[8]を用いた。また、3.3.4項の各重みを、 $W_L = 0.6, W_H = 0.2, W_C = 0.2$ と設定した。

4.2 評価指標:BLEU

正解文と提案した手法により生成された文の評価にはBLEUを用いた($N = 4$)。BLEUとは、用意された参照文と機械翻訳システムが出力した文とのn-gramマッチ率を測る指標である。レシピ文自動生成において、(1)文法、(2)文の曖昧性、(3)意味の正確さ、に対する人手による評価と、BLEUのスコアに相関が認められる[3]ので、本研究でも評価指標としてBLEUを採用した。なお、本研究では1レシピをBLEUの単位として扱った。

4.3 結果

(1) 頻度のみ(F),(2) 内容語残存コストのみ(C),(3) エントロピーのみ(E),(4) 頻度+内容語残存コスト(F+C),(5) 頻度+エントロピー(F+E),(6) 内容語残存コスト+エントロピー(C+E),(6) 頻度+内容語コスト+エントロピー(F+C+E)それぞれの手法において、テンプレート生成時の学習コーパスを変化させた時のBLEUの値を表3に示す。

4.4 考察

先行研究[3]ではテンプレートの頻度のみ(表3中F)に着目した手法であったが、本研究でエントロピーと内容語残存コストの導入で精度が向上することを示した。特にF+E及びF+C+Eが比較的高い水準の結果となった。これは、エントロピーの導入によって活用語尾や適切な助詞

表3 コーパスサイズに対する各手法のBLEU

コーパスサイズ 手法\	18,590 (1/64)	74,362 (1/16)	297,448 (1/4)	1,189,794 (1/1)
F	8.78	12.65	11.82	13.05
C	9.23	12.30	10.85	12.72
E	11.93	12.10	11.85	11.18
F+C	10.11	12.21	11.92	12.78
F+E	10.45	16.05	16.93	17.75
C+E	12.37	16.89	15.99	14.80
F+C+E	10.79	15.93	17.45	17.80

の接続がなされるようになり、BLEUの評価であるn-gramのマッチ率に対する影響力が大きいからだと考えられる。また、本研究では頻度、内容語残存コスト、エントロピーをテンプレートの尤度計算に用いたが、それぞれ単独による評価では頻度(F)による評価が最も高く、先行研究[3]の有用性を示した。

EからE+CはBLEUの改善が見られるが、FからF+C、F+EからF+C+Eと内容語残存コストを追加した場合、あまりBLEUの変化は見られない。これは、BLEUはあくまでも原文とのn-gramマッチ率を見ているのであり、局所的な文法の正しさを反映することに起因すると考えられる。実際、F+EとF+C+Eで出力された40レシピの生成文を実際に調理を行うという視点で比較したところ、F+Eでは原文にはない材料や形容語を含むテンプレートを選択しているレシピが散見したが、F+C+Eでは明らかにその頻度が減っていることが確認できた。

本研究では、3.3.4項:尤度計算で、尤度に対する各重み W_L, W_H, W_C を特定の値で求めている。しかし、それらの重みはグリッドサーチ等によって最適な値を求めべきである。

5. おわりに

本研究ではフローグラフからレシピ文を自動生成する手法について述べた。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尤度計算によってテンプレートを選択する。尤度計算の手法として、テンプレート頻度、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。429,691レシピから学習を行い、40レシピを入力として与え、生成された文とレシピの原文との比較をBLEUで評価し、精度向上に寄与することを確認した。内容語残存コストを導入することで、生成された文の中から内容語が減少することを目視で確認した。

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, 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, *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, pp. 2370–2377 (2014).
 - [3] Mori, S., Maeta, H., Sasada, T., Yoshino, K., Hashimoto, A., Funatomi, T. and Yamakata, Y.: FlowGraph2Text: Automatic Sentence Skeleton Compilation for Procedural Text Generation, *Proceedings of the the Eighth International Conference on Natural Language Generation*, pp. 118–122 (2014).
 - [4] Yamakata, Y., Imahori, S., Sugiyama, Y., Mori, S. and Tanaka, K.: Feature Extraction and Summarization of Recipes using Flow Graph, *Proceedings of the 5th International Conference on Social Informatics*, LNCS 8238, pp. 241–254 (2013).
 - [5] Naim, I., Song, Y. C., Liu, Q., Kautz, H., Luo, J. and Gildea, D.: Unsupervised Alignment of Natural Language Instructions with Video Segments, *Proceedings of the 28th National Conference on Artificial Intelligence* (2014).
 - [6] Rohrbach, M., Qiu, W., Titov, I., Thater, S., Pinkal, M. and Schiele, B.: Translating Video Content to Natural Language Descriptions, *Proceedings of the 14th International Conference on Computer Vision* (2013).
 - [7] Yu, H. and Siskind, J. M.: Grounded Language Learning from Video Described with Sentences, *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (2013).
 - [8] Papineni, K., Roukos, S., Ward, T. and Zhu, W.-J.: BLEU: a Method for Automatic Evaluation of Machine Translation, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 311–318 (2002).
 - [9] 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).
 - [10] Mori, S., Sasada, T., Yamakata, Y. and Yoshino, K.: A Machine Learning Approach to Recipe Text Processing, *Proceedings of the 1st Cooking with Computer Workshop*, pp. 29–34 (2012).
 - [11] 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).