

手続きの生成モデルの文法抽出の枠組みによる学習

吉成未菜里[▲] 横井祥^{▲◇} 乾健太郎^{▲◇}

[▲] 東北大学 [◇] 理化学研究所 AIP センター

{yoshinari, yokoi, inui}@ecei.tohoku.ac.jp

1 はじめに

料理の作り方, 各種機器のマニュアル, 実験で収集したデータの解析の流れなど, 事物を利用した一連の手順を表した「手続き」に対して, 検索, 要約, クラスタリングなどの各種情報処理が盛んに研究されている. 手続きの中でも, 特に料理の手順を示したレシピはデータが潤沢なこともあり盛んに研究されている [9, 3, 5, 6]. 本研究でも手続きの一例としてレシピを対象とする.

手続きの研究は大きく 2 種類に分けられる. 1 つ目は, 自然言語文で書かれた手続きを木やグラフに構造化する研究である. Mori らはレシピの無閉路有向グラフによる意味表現を定義し, 日本語レシピ投稿サイトであるクックパッドのレシピの一部を用いて意味理解のためのコーパスを作成した [3]. また Tasse らは, レシピの一階述語論理を基にした意味表現 MILK を定義し, MILK によって英語のレシピをアノテーションすることでコーパスを作成した [5].

2 つ目は, 木やグラフに構造化された手続きに対して要約, 汎化, 検索などの情報処理をおこなう研究である. 山肩ら, レシピの典型性を評価するため, 木構造で表現されたレシピの間の編集距離を算出するアルゴリズムを提案した [8]. Wang らは, 調理手順の類似性に基づいたレシピ検索を実現するため, レシピ間の類似度を部分グラフの頻度を用いて定義した [6]. 事例ベース推論の文脈においては, Müller らがドメインのオントロジー知識を利用してワークフローの一般化・詳細化を試みている [4]. 本研究も, あらかじめ木に構造化されている手続きに対して情報処理を試みる.

本研究の提案は, (1) 木構造で表現された手続きに対してその確率的生成モデルを学習すること, また (2) これを文法抽出の枠組みで実現すること, の 2 点である.

まず, 手続きに対する確率的生成モデルを学習すれば, これまで個別に解かれてきた既存タスクを統一的に解決できる可能性がある. 例えば, 山肩ら [8] のおこなったレシピの典型性の評価は, 当該品目の多数のレシピから学習した「レシピの生成モデル」における当該レシピの生成確率を典型性の度合いと見なすことで実現できる. あるいは, Wang ら [6] のおこなった類似レシピ検索は, あるレシピの品目 (例えばシチュー) の生成モデルにおける他のレシピ (例えばカレーのレシピ) の生成確率として品目間の類似度を定義すれば実現できる. また, Müller らが指摘しているように, 手続きとして書かれた知識を新しい状況に適用するためには, 既知の手続きに対する一般化が必要である [4]. 確率的生成モデルの学習はデータの集合に対する統計的な一般化であり, Müller らが取り組んだような事例ベース推論への応用も期待される.

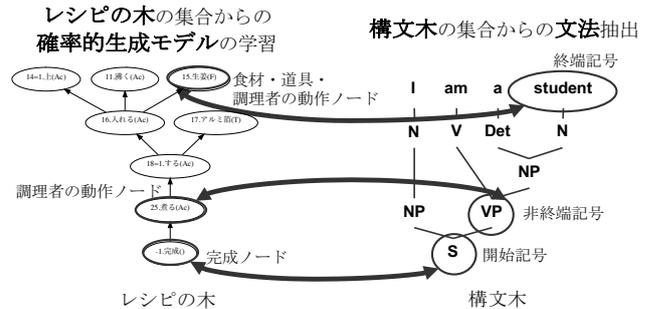


図 1: 「手続きの木の集合からの確率的生成モデルの学習」と「構文木の集合からの文法抽出」の対応

また, 木で表現された手続きに対する確率的生成モデルの学習は, 構文木の集合から文法を学習する問題と捉えることができる (図 1, 3.1 節). このような問題は, 自然言語処理の分野では文法抽出のタスクとして広く研究されており [1], 様々な文法や学習法が提案されている. 本稿では文法クラスとして, 最も標準的な確率的文法である確率文脈自由文法 (PCFG) を採用する (3.2 節).

実験では, レシピの実データを用いて品目 (例えば肉じゃが) ごとの PCFG を学習し, 提案手法によってその品目らしさ (例えば肉じゃがらしさ) を捉えられるかを確認した. 類似レシピ検索への適用では提案手法の有用性が確認できたが, 典型的レシピの生成では PCFG の表現力が限定的であることに起因する課題が見つかった.

以上のように本稿では, 手続きに対する確率的生成モデルを文法抽出の枠組みで学習することで, 手続きに対する各種情報処理を統一的に解決することを試みる. また文法として特に PCFG を用いた実験を行い, その有用性や課題について論じる.

2 データ・問題設定

2.1 データ

以下本稿では, 図 2 に示す木構造を持ったレシピデータを扱う. 木のノードには「食材ノード (F)」「道具ノード (T)」「調理者の動作ノード (Ac)」の 3 種類がある. 食材と道具に対して動作 (調理) を繰り返し, 完成品である根ノードができあがる.

本研究で扱ったレシピの木は合計 70 個で, その内訳は, 「サバの味噌煮」「ビーフシチュー」「豚のしょうが焼き」「ハンバーグ」「肉じゃが」「ポテトコロッケ」「ロールキャベツ」の 7 品目それぞれにつき 10 件である*1. 以下でデータの入手方法と

*1 最大の日本語レシピ投稿サイトであるクックパッド

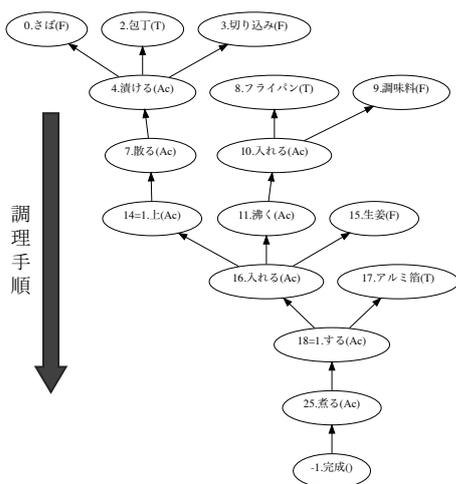


図 2: サバの味噌煮のレシピの木構造の例

木構造への変換手順を示す。

レシピフローグラフで表現されたデータの入手

はじめに山肩ら [8] から、当該 70 件のレシピ (生テキスト) を **レシピフローグラフ** に変換したデータの提供を受けた*2。レシピフローグラフは Mori らが提案したレシピの抽象表現で、レシピ中に登場する用語 (食材名「人参」、動作名「切る」など) をノードとし、それらの間の関係 (「切る」という動作で「人参」を調理する、など) をエッジとした有向無閉路グラフ (DAG) である [3]。生文のレシピをレシピフローグラフに変換するパーザも開発されており、生文からグラフ構造への変換の F 値で 78.3 という性能が出ている [10]。本稿では小規模データセットに対する実験をおこなったが、今後パーザを用いてデータサイズを拡大することは容易である。

前処理と木構造への変換

レシピフローグラフ表現を以下の方法で木構造に変換した。

- **「する」の削除**: 生文における「みじん切りにする」は、レシピフローグラフ表現では「みじん切り」と「する」のふたつに分割される。「する」は動作ノードとしてほとんど意味を持たないので、「みじん切り」のみを残して「する」を削除する。
- **同義語のまとめ上げ**: 今回利用したデータは 7 品目各 10 件の計 70 件のレシピからなる小規模なデータであるため、データのスパース性を事前にできるだけ緩和しておく必要がある。土居らの料理オントロジ [11] を用いて、動詞と名詞について同義語をまとめ上げる。
- **木構造への変換**: Yamakata らの手法 [7] を参考にノードとエッジを削除し、レシピフローグラフを「食材」「道具」「調理者の動作」からなる木に変換する。

2.2 問題設定

本研究で取り組む問題は以下の通り (図 3):

- **入力**: 同一品目のレシピの木の集合

$$D = \{t_1, \dots, t_n\} \subseteq \mathcal{T} \quad (1)$$

(<http://cookpad.com/>) で、各品目名をクエリとしたときの人気順上位各 10 件を、後述の方法で木構造に変換した。

*2 我々が扱うレシピ 70 件は山肩らの研究 [8] で扱われているものと同じであるが、前処理の方法 (木構造への変換方法) に若干の違いがある。

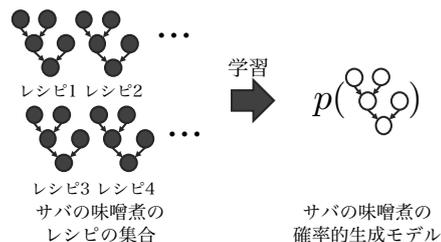


図 3: 手続きの木の集合から確率的生成モデルを学習する例

(例えばサバの味噌煮のレシピの木の集合),

- **出力**: D の対数尤度

$$\ell(D) = \log \prod_{i=1}^n p(t_i) = \sum_{i=1}^n \log p(t_i) \quad (2)$$

を最大化するような確率的生成モデル p (例えばサバの味噌煮のレシピの確率的生成モデル)。

3 提案手法

本研究では、手続きの木の集合からその手続きの確率的生成モデルを学習する問題を、文法抽出の枠組みで解く。また、文法クラスとして標準的な確率文法である確率文脈自由文法 (PCFG) を考える。すなわち、PCFG の学習を通して手続きの生成モデルの学習を試みる。

3.1 木の生成モデルの学習と文法抽出の関係

手続き (調理レシピ) の木の集合からの確率的生成モデルの学習が、構文木の集合からの文法抽出だと捉えられる (図 1) ことを述べる。文法抽出と手続きの木の生成モデルの学習の各要素の対応は図 1 のとおり。本研究のタスクにおける手続きの木は、文法抽出における構文木に相当する。また、レシピの木と構文木のノードは以下のように対応が取れる。

- 完成品ノード \leftrightarrow 開始記号
- 調理者の動作ノード \leftrightarrow 非終端記号
- 食材ノード, 道具ノード \leftrightarrow 終端記号

このように、レシピの木の集合を構文木の集合と見れば、レシピの木の生成モデルは文法抽出の枠組みで学習できることが分かる。

ただし、レシピの木と構文木には、非終端記号の解釈に大きな違いがある。レシピの木における非終端記号すなわち調理者の動作ノード (例えば「煮る」) は、生文中で観測できる語彙項目である。一方構文木における非終端記号 (例えば VP) は、生文で観測できる語彙項目ではない。では構文木における非終端記号 (例えば VP) は何かと言えば、その子孫ノードにあたる終端記号たち (例えば “have a pen”) の構成結果である。一方で、レシピの木における非終端記号 (例えば「煮る」) は、子孫ノードの集合の構成結果 (例えば「肉と野菜に下ごしらえをして煮たもの」とは異なる。「調理者の動作ノード」は構成の一番最後に最後に適用された関数を表す記号に過ぎない。このようにレシピの木と構文木は完全に対応するわけではないが、本研究では十分な対応関係があると考え、文法抽出の枠組みで木の生成モデルの学習を試みる。

3.2 モデル: 確率文脈自由文法 (PCFG)

本稿では文法クラスとして、標準的な確率文法である確率文脈自由文法 (PCFG) を採用する。

学習 生成規則の学習は最尤推定による。学習データに含まれない記号（食材、道具など）に対処するためラプラス・スムージングによる平滑化をおこなう。また、一般的な PCFG の学習と異なり、生成規則の右辺の記号列の並び順を無視する。今回の問題設定では右辺の記号列の順番の違いは意味をなさない。例えば「煮る → じゃがいも, 人参」と「煮る → 人参, じゃがいも」は同一の生成規則と考えるのが自然である。

生成 実験では、学習済みの PCFG（例えば「肉じゃが文法」）から最も対数尤度の高い手続きの木（例えば典型的な肉じゃがのレシピ）を生成することを試みた。大域的に対数尤度が最大となる手続きを生成するには、導出（開始記号から適用できる生成規則の系列）の全体を探索空間とする最適化が必要であり、これは現実的ではない。そこで本稿では、ビーム幅 20、木全体のノード数が 80 になったら打ち切りという条件でビームサーチをおこない、この範囲で最も対数尤度の高い木を解とした。

4 実験

手続きの木の確率的生成モデルを、文法抽出の枠組みで、特に PCFG の推定の枠組みで学習することの妥当性を確認するため、2.1 節に示した 7 品目、各 10 件の料理レシピに提案手法を適用した。文法（確率的生成モデル）は同一品目のレシピの集合に対して学習した。例えば、肉じゃがのレシピの集合から「肉じゃが文法」を学習した。

はじめに学習された PCFG の生成規則の一部を示し、品目毎に学習した文法がその品目らしさ（例えば肉じゃがらしさ）を捉えられているか確認する（4.1 節）。また、最も尤度の高いレシピの木の生成結果を確認する（4.2 節）。次に情報検索の枠組みで提案手法を評価する。品目毎に学習した文法（例えば「肉じゃが文法」）が、テストセットの中で当該品目のレシピ（例えば肉じゃがのレシピ）をどの程度高くランキングするかを確かめる（4.3 節）。

4.1 定性的評価: 学習された生成規則の確認

PCFG で「品目らしさ」を学習できるかを確かめるため、7 品目（肉じゃが、サバの味噌煮、…）それぞれの文法 (PCFG) を、各々 10 件すべてのレシピデータを使って学習し、実際に学習された生成規則を確認した。表 1 に「豚のしょうが焼き文法」における生成規則を示す。表の中には、豚のしょうが焼きの調理で見られる典型的な（調理動作、材料）の組

- (13) 豚肉を切る
 - (15) 玉ねぎをスライスする
- や、調理一般に見られる典型的な（調理動作、道具）の組
- (12) キッチンペーパーで拭く
 - (14) 電子レンジでチンする

が学習できていることが確認できた。

4.2 定性的評価: 典型的なレシピの生成

4.1 節と同様の設定で品目毎の文法を学習した上で、学習された品目毎の文法から最尤のレシピを生成し、当該品目らしい（典型的な）レシピを生成できるか確認した。実際に「ハンバーグ文法」と「ビーフシチュー文法」から生成したレシピを図 4 に示す。

図 4 の木を見ると、いくつかの類似した問題が確認できる：

表 1: 豚のしょうが焼きのモデルにおける生成規則

生成規則	確率値
(1) 完成 → 絡める (Ac)	0.16
(2) 完成 → かける (Ac)	0.16
(3) かける (Ac) → ごま (F), 盛り付ける (Ac)	0.29
(4) かける (Ac) → 煮詰める (Ac), 盛り付ける (ac)	0.29
(5) かける (Ac) → つける (Ac), 付け込む (Ac), 合わせる (Ac)	0.29
(6) つける (Ac) → 入れる (Ac), 炒める (Ac), 飲料水 (F)	0.29
(7) つける (Ac) → 加熱 (Ac), 焦がさない (Ac)	0.29
(8) つける (Ac) → 裏返す (Ac)	0.29
(9) 裏返す (Ac) → 並べる (Ac), 付け込む (Ac), 焼き色 (F)	0.67
(10) 並べる (Ac) → する (Ac), 付け込む (Ac)	0.40
(11) 並べる (Ac) → 拭く (Ac), 熱す (Ac)	0.40
(12) 拭く (Ac) → キッチンペーパー (T), 切る (Ac)	0.67
(13) 切る (Ac) → 豚肉 (F)	0.22
⋮	⋮
(14) チン (Ac) → スライスする (Ac), 電子レンジ (T)	0.67
(15) スライスする (Ac) → 玉葱 (F)	0.75
⋮	⋮

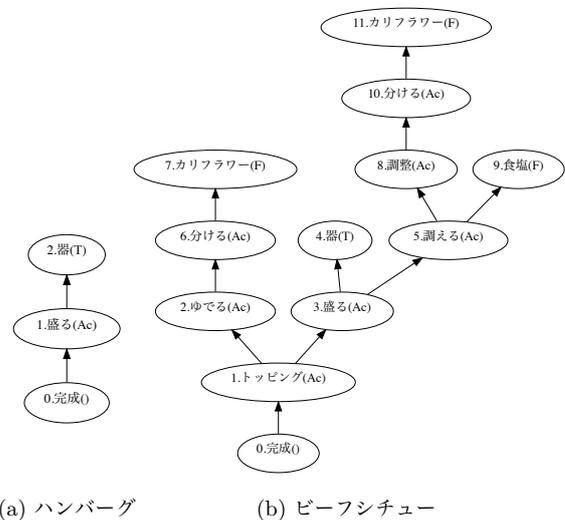


図 4: 各品目のモデルから生成したレシピの木

- 図 4a では、「肉」などのハンバーグに必要な食材が（1 つも）使われていない。
- 一方で、図 4b ではカリフラワーを 2 回使っている。
- 図 4a では、「器に盛る」という手順を生成しているにも関わらず、盛られる料理の食材や、調理のための動作ノードが一切登場しない。

すなわち「必要な材料を網羅しているか」「必要な調理動作を全て含んでいるか」といった情報を捉えるのに提案手法では限界があると考えられる。

同様の問題は再帰型ニューラルネットワーク (RNN) によってレシピの生文を生成する研究でも指摘されている [2]。Kiddon らは、Neural Checklist Model を提案し、「これまでどんな内容を書いたか、これからどんな内容を書く必要があるのか」の制御を試みた。本研究においても、Kiddon らが導入したようなチェックリストの導入によって材料の管理の問題を解決できる可能性がある。

4.3 定量的評価: 類似レシピ検索

品目ごとに学習した生成モデル（「肉じゃが文法」、「サバの味噌煮文法」、…）のすべてでテストセット（「肉じゃが」、「サ

表 2: モデル・検索対象ごとの MRR (太字は検索対象ごとの最も高い MRR を示したモデル)

モデル \ 検索対象	ビーフシチュー	豚のしょうが焼き	ハンバーグ	肉じゃが	ポテトコロッケ	ロールキャベツ	サバの味噌煮
ビーフシチュー	0.282	0.633	0.235	0.507	0.260	0.193	0.483
豚のしょうが焼き	0.221	0.733	0.239	0.468	0.232	0.210	0.489
ハンバーグ	0.223	0.600	0.356	0.398	0.221	0.222	0.573
肉じゃが	0.279	0.633	0.230	0.492	0.245	0.230	0.483
ポテトコロッケ	0.258	0.542	0.338	0.465	0.371	0.202	0.418
ロールキャベツ	0.196	0.567	0.264	0.385	0.270	0.402	0.509
サバの味噌煮	0.235	0.508	0.285	0.482	0.217	0.232	0.633

バの味噌煮, ...) を対数尤度によりランキングしたとき, 「サバの味噌煮」に最も高い順位をつけたのが「サバの味噌煮文法」であれば, 文法はその品目らしさを表していると言えよう. 我々はこの直感の通り,

- **検索クエリ**: 各品目毎に 9 件のレシピで学習した生成モデル {「肉じゃが文法」, 「サバの味噌煮文法」, ...}
- **検索対象**: 各レシピが 1 件ずつ入った系 7 件のテストセット {「肉じゃが」, 「サバの味噌煮」, ...}

を用意し, (検索クエリ, 検索対象) の組合せ毎に順位を計測した. これを leave-one-out 交差検証の要領で 10 回繰り返して, その MRR (平均逆順位) を報告する.

4.3.1 実験結果

実験計算結果を表 2 に示す. 表の縦一列の中で灰色のセルの数値が一番高ければ (太字であれば), 検索対象のレシピを最も高くランキングしたのは当該品目の文法, ということになる. 例えば一番左の列を見ると, ビーフシチューを最も高くランキングしているのは「ビーフシチュー文法」であることが分かる. 肉じゃが以外のすべての検索対象レシピでこのような良い性質が示された. 肉じゃがだけは, 最も MRR が高いモデルは「ビーフシチュー文法」で, 次点が「肉じゃが文法」となっている.

また, 表 2 をモデルの品目ごと (横) に見ていくと, 「ビーフシチュー文法」, 「ハンバーグ文法」, 「ポテトコロッケ文法」, 「ロールキャベツ文法」については同じ品目の検索対象の MRR が低くなっていることが分かる.

4.3.2 考察

表 2 をモデルの品目ごと (横) に見ると, モデルの品目と同じ品目の検索対象の MRR が低くなる品目 (ビーフシチューなど) が存在する. これは PCFG において, ノード数の大きな木の対数尤度が低くなりやすいためであると考えられる. PCFG における木の対数尤度は, その木を生成する際に適用される生成規則の確率値の積の対数をとったものである. ノード数の大きな木はそれだけ適用される生成規則が多くなりやすいため, 対数尤度が低くなりやすい. 例えば, 本実験の対象である 7 品目ではポテトコロッケの木の平均ノード数が最も大きい, 表 2 でもポテトコロッケの MRR は「ポテトコロッケ文法」を含むどのモデルにおいても低い.

また, 肉じゃがに最も高い MRR を与えるモデルは「ビーフシチュー文法」であるが, ビーフシチューと肉じゃがはいずれも肉を野菜と一緒に煮込む料理であり, 似た調理手順をもつ品目であるといえる. このように提案手法による尤度計算は, 品目間の調理手順の類似度を反映すると考えられる.

5 おわりに

本稿では, 手続きに対する確率的生成モデルを学習すること, またこれを文法抽出の枠組みで実現することを提案した. またこれらの提案に対し, 文法として PCFG, 手続きの代表例としてレシピデータを用いた評価実験をおこなった. 実験の結果, 類似レシピ検索では提案手法の有用性が確認できたが, 典型的レシピの生成では適切なレシピの木を生成することができず, 子孫ノードを生成するのに必要な情報を保持できないなどの課題が見つかった.

今後この問題を解決するため, 手続きを表現するためのモデルをより適切なモデルに改良したい. 候補の 1 つとして, lexicalized PCFG が考えられる. lexicalized PCFG は本稿で用いた PCFG の拡張で, PCFG の各生成規則に語彙の情報を含めたものである. lexicalized PCFG を用いれば, 子孫ノードを生成するために必要な情報を保持できないという提案手法の問題点に対処できると考えられる. 手続きの特徴を十分に表現可能なモデルを作成した上で, 手続きの確率的生成モデルを各タスクへ適用することが将来的な課題である.

参考文献

- [1] Eugene Charniak. Tree-bank Grammars. In *Proceedings of the National Conference on Artificial Intelligence*, pp. 1031—1036, 1996.
- [2] Chloé Kiddon, Luke Zettlemoyer, and Yejin Choi. Globally Coherent Text Generation with Neural Checklist Models. In *{Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing}*, pp. 329—339, 2016.
- [3] Shinsuke Mori, Hirokuni Maeta, Yoko Yamakata, and Tetsuro Sasada. Flow Graph Corpus from Recipe Texts. In *The 9th edition of the Language Resources and Evaluation Conference*, pp. 2370—2377, 2014.
- [4] Gilbert Müller and Ralph Bergmann. Generalization of Workflows in Process-Oriented Case-Based Reasoning. *The Twenty-Eighth International Flairs Conference*, pp. 391—396, 2015.
- [5] Dan Tasse and Noah A Smith. Toward semantic processing of recipes. Technical Report CMU-LTI-08-005, 2008.
- [6] Liping Wang, Qing Li, Na Li, Guozhu Dong, and Yu Yang. Substructure similarity measurement in chinese recipes. *Proceeding of WWW '08*, pp. 979—988, 2008.
- [7] Yoko Yamakata, Shinji Imahori, Hirokuni Maeta, and Shinsuke Mori. A method for extracting major workflow composed of ingredients, tools, and actions from cooking procedural text. *2016 IEEE International Conference on Multimedia and Expo Workshop, ICMEW 2016*, 2016.
- [8] 山岡洋子, 今堀慎治, 森信介, 田中克己. ワークフロー表現を用いたレシピの典型性評価と典型的なレシピ表現. *電子情報通信学会論文誌, Vol. J99-D, No. 4*, pp. 378—391, 2016.
- [9] 山岡洋子, 森信介. ユーザ投稿型レシピの情報処理. *情報処理, Vol. 57, No. 4*, pp. 374—382, 2016.
- [10] 前田浩邦, 山岡洋子, 森信介. 手順文書からの意味構造抽出. *人工知能学会論文誌, Vol. 32, No. 1*, pp. E-G24.1—8, 2017.
- [11] 土居洋子, 辻田美穂, 難波英嗣, 竹澤寿幸, 角谷和俊. 料理レシピと特許データベースからの料理オントロジーの構築. *電子情報通信学会技術研究報告 = IEICE technical report: 信学技報, Vol. 113, No. 469*, pp. 37—42, 2014.