

zero-shot 学習を用いた言語資源の転移に関する一考察

樺山絵里 † 小林一郎 † 麻生英樹 ‡ 持橋大地 ¶ 中村友昭 § 長井隆行 §
 †お茶の水女子大学 ‡産業総合技術研究所 ¶統計数理研究所 §電気通信大学

1 はじめに

本研究では、人の動作を説明する自然言語文から、バイグラムモデルを構築し、動的計画法を用いて、尤度の高い学習の単語の組み合わせを獲得することによりテキスト生成を行う。とくに先行研究 [1][2] で提案された手法の汎用性を検証するため、テキスト生成のための言語資源の転移について考察を行う。

2 言語モデルの推定における zero-shot 学習

人の動作を記述する自然言語表現を考えた際に、図 1 に示すように対象とする動作を「上げる/下げる」「速く/普通/遅く」「高く/低く」「前から/横から」「手/足」「左/右/両」の組み合わせの 120 の動作から構成されると定義する。また、1 動作あたり上記の 6 記述要素が含まれるとする。例えば、「左手を前から高く速く上げる」の場合、「左」「手」「前から」「高く」「上げる」という要素から成ると考えることができる。また、本研究では、120 の動作に対する自然言語記述を収集し、そこから構築したバイグラムモデルを用いて、尤度が高くなるような単語の組み合わせを見つけることにより文の生成を行う。文長に左右されないテキスト生成を行うために、小林ら [1] が用いた、疑似単語 (番号付き null) をバイグラムモデルに導入することにより文長に関わらず尤度の次数を同じにするテキスト生成手法を適用する (図 2 参照)。

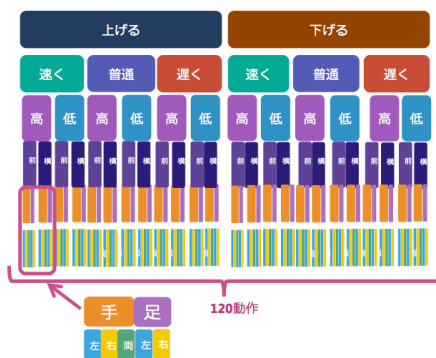


図 1: 動作カテゴリの構成



図 2: 仮定の単語 null の導入

3 zero-shot 学習に基づく言語資源推定

特定の意味を説明するための言語資源が存在しないことの影響を評価するために、一部の動作に対する言語資源を取り除き、最小二乗推定による zero-shot 学習を行い、他の動作に対する言語資源用いて、欠損している動作に対する言語モデルを推定する。その後、推定された言語モデルを用いて説明文の生成を行い、得られた説明文の品質を評価する。

zero-shot 学習は、動作カテゴリに対して、バイグラムを表した行列 Ψ を、それが含んでいる動作要素を表した行列 A およびそれぞれの動作要素に対して使用される言語資源の確率を表す行列 Φ とに分解できるということを前提に、特定の動作カテゴリのバイグラムが欠損している場合、行列の一般逆行列を使って欠損部分の要素を推定する。上述した内容は式 (1) で表せる。ここで、 A^+ は A の一般逆行列を示す。

$$\hat{\Phi} = \min_{\Phi} \|\Psi - A\Phi\|^2 = A^+\Psi \quad (1)$$

4 実験

図 1 に示すものを実験データとして説明する。学習用データが存在しないことの影響を評価するために、データセットの一部の動作に対するデータを取り除き、最小二乗推定による zero-shot 学習を行うことにより、他の動作に対するデータを用いて、データ欠損行動に対する言語モデルの推定を行う。その後、推定された言語モデルを用いて説明文の生成を行い、得られた説明文の品質を評価する。

4.1 実験設定

zero-shot 学習により、データ欠損行動の言語モデルをどの程度正確に推定可能であるかを検証するために、動作のカテゴリ構成においてランダムに取り除くようにして、以下の 4 つの場合について検討した。各場合において 10 回ずつ実験を行いその平均をとる。

1. full (言語資源を全て使用)
2. three-quarters (4 分の 3 を使用)
3. half (半分を使用)

A study on Transference of Linguistic Resources using zero-shot Learning

†Eri KABAYAMA(kabayama.eri@is.ocha.ac.jp)

†Ichiro KOBAYASHI(koba@is.ocha.ac.jp)

‡Hideki Asoh ¶Daichi Mochihashi §Takayuki Nagai

†Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University

¶National Institute of Advanced Industrial Science and Technology

§The Institute of Statistical Mathematics

§The University of Electro-Communications

表 1: 「速く大きく左足を横から上げる」という動作に対する削減された言語資源の下での生成文

言語資源	生成文
full	<ul style="list-style-type: none"> •左足を素早く大きく横から上げる。 null8 null9 null10 •左足を大きく横から上げる。 null8 null9 null10 EOS •さっと左足を素早く大きく横から上げる。 null8 null9
three-quarters	<ul style="list-style-type: none"> •左足を大きく横から上げる。 null8 null9 null10 EOS •左足を素早く大きく横から上げる。 null8 null9 null10 •さっと左足を素早く大きく横から上げる。 null8 null9
half	<ul style="list-style-type: none"> •左足を大きく横から上げる。 null8 null9 null10 EOS •左足を素早く大きく横から上げる。 null8 null9 null10 •さっと左足を素早く大きく横から上げる。 null8 null9
a-quarter	<ul style="list-style-type: none"> •左足を大きく横から上げる。 null8 null9 null10 EOS •左足を素早く大きく横から上げる。 null8 null9 null10 •さっと左足を素早く大きく横から上げる。 null8 null9

表 2: BLEU スコアおよび生成文の対数尤度に基づく評価結果

	推定された言語資源	full	three-quarters	half	a-quarter
BLEU	(データ全動作)	1.0	0.9258	0.8716	0.7847
	欠損動作	(1.0)	0.7415	0.7648	0.7130
対数尤度	a-quarter, half, three-quarters 共通欠損動作	-6.019	-60.76	-60.77	-60.76
	half, three-quarters 共通欠損動作	-2.999	-60.93	-60.93	—

4. a-quarter (4分の1を使用)

生成された文の定量的な評価手法として、以下の2つを考える。

- BLEU スコアによる評価

full のデータから学習した言語モデルによって生成されたテキストと zero-shot 学習によって推定された言語モデルによって作成されたテキストとの BLEU スコアにより評価する。

- 生成文の尤度評価

zero-shot 学習によって推定された言語モデルから生成された尤度が上位3件の説明文の対数尤度を full のデータから学習した言語モデルを用いて算出した際の平均値によって評価する。このとき、full の言語モデルの中に推定された言語モデルから生成された文に現れる単語ペアがない場合には、線形補間法を用いた。

4.2 実験結果

言語モデル full, three-quarters, half, a-quarter に対して、それら全てに共通して推定された言語モデルである動作「速く大きく左足を横から上げる」に関するテキスト生成結果を表1に示す。

4.3 評価結果

4.3.1 BLEU スコアによる評価

表2に zero-shot 学習によって推定された言語モデルおよび取り除く対象にならなかった言語モデルの両方を用いて、全動作に対するテキストを評価した結果を示す。また、欠損動作のみに対しても評価を行った。データ全動作の評価結果は、言語資源が豊富なほど値が高くなった。欠損動作の評価結果は、言語資源の量によらない結果となった。

4.3.2 生成文の性能評価

three-quarters, half, a-quarter について共通するデータ欠損動作に対して、zero-shot 学習で推定された言語モデルから生成された文の対数尤度を、full の言語モデルで計算した。full, three-quarters, half の3つのケースをより詳しく比較するため、three-quarters と half に共通するデータ欠損動作についての評価も実施した。表2にその結果を示す。full の値が高くなっており、three-quarters, half, a-quarter の間での差は見られなかった。

4.4 考察

推定した言語モデルおよびそれらのモデルから生成した文は、言語資源の量にあまりよらない。本研究で用いたデータセットの文数が少なかったために、全体的に精度が低くなったと考える。

5 おわりに

本研究では、バイグラムに基いてテキストを生成する際に十分な言語資源が存在しない場合、zero-shot 学習によってどの程度の資源の拡充が可能化を考察した。今後の課題として、データセットの拡充を行ない、より多くのデータの下、提案手法における言語資源転移の性能評価を行ないたい。

参考文献

- [1] 小林瑞季, 麻生英樹, 小林一郎, 人の動作を対象にした確率的言語生成への取り組み, 言語処理学会第20回年次大会, pp.920-923, 北海道大学, 2014.
- [2] 樺山絵里, 小林一郎, 麻生英樹, 持橋大地, アッタミムハンド, 中村友昭, 長井隆行, 人の行動を表現するテキスト生成の一考察, 第30回人工知能学会全国大会, 3P1-2in2, 北九州国際会議場, 2016.