

# RNNによる英文空所補充問題の解法の検討

Examination of solution of cloze deletion tests for English sentences by recurrent neural network

玉城 悠仁

Haruhito Tamaki

岡山大学 太田研究室

Ohta Laboratory, Okayama University

**概要** 近年、ニューラルネットワークを用いた自然言語処理の研究が盛んに行われており、特に sequence-to-sequence (seq2seq) モデルは機械翻訳や文章要約で高い成果をあげている。本研究では seq2seq モデルを用いて英文空所補充問題の解答に取り組む。大学入試センター試験の空所補充問題について実験を行い、その正答率を評価した。

## 1 はじめに

現在、日本人の多くが、第二言語として英語を学んでいる。しかし、英語を母語としない日本人にとって、正しい英語の使い分けは難しく、日本人の英作文には様々な誤りがしばしば見られる。

そういった学習者を支援するために英文誤り訂正の研究が行われているが、近年はニューラルネットワーク (neural network; NN) を用いた手法が注目されている。機械翻訳や文章要約において sequence-to-sequence (seq2seq) モデル [1] [2] が高い成果をあげており、文法誤り訂正のモデルとしても提案されている。seq2seq モデルではしばしば attention メカニズム [3] [4] が使われており、これにより系列の長さに関わらず、入力系列の情報を出力系列へと反映することが可能となる。Xie ら [5] は attention に基づく seq2seq モデルを用いて英文誤り訂正を行い、代表的なベンチマークデータセットである CoNLL-2014 Shared Task データセットにおいて高い精度を達成できることを示した。

本研究では、attention に基づく seq2seq モデルの英文の空所補充問題への適用を試み、その性能を確認する。

## 2 seq2seq モデルによる英文空所補充

本研究では attention に基づく seq2seq モデルを用いた英文空所補充問題の解答に取り組む。本研究では Luong ら [4] の seq2seq のモデルの構造、パラメータ、及び最適化アルゴリズムについて同じものを用いる。

提案手法の概略を図 1 に示す。図 1 では空所を含む英文 “I am a { }” が入力系列、空所を補充した英文 “I am a { student }” が出力系列の際の、attention の算出の最初のステップを表している。“<s>” はデコードの開始位置、“</s>” はデコードの終了位置を表す。Luong らのモデルは英仏翻訳のモデルであったが、

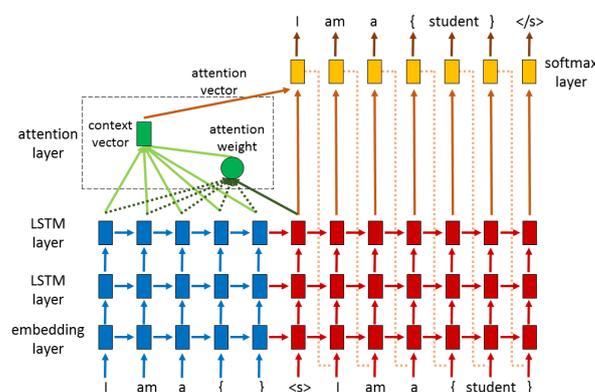


図 1: 実験で用いるモデルの概要

“{” を空所の開始位置，“}” を空所の終了位置を表す単語として扱うことで、英文空所補充問題に seq2seq の考え方を適用できるようにした。

seq2seq モデルは encoder-decoder モデルとも呼ばれ、encoder 部と decoder 部から成る。本モデルでは encoder 部、decoder 部ともに 128 次元の 2 層の LSTM である。本モデルの出力は attention とデコード結果を softmax による処理を行うことで生成された、単語の確率分布である。また、損失関数はモデルの出力と教師データとの交差エントロピー誤差であり、この損失が最小となるように学習を行う。

## 3 評価実験

### 3.1 実験データ

学習データおよび検証データには text8 コーパス\*1を用いる。text8 コーパスを無作為に分割し、分割後の一部を空所に置換した英単語列 50000 件を学習データ、1000 件を検証データとした。

テストデータには 2000 年度から 2016 年度に行われた大学入試センター試験の本試験および追試験の大問 2A のうち、空所が 1 箇所のみである 324 問を使用した。大学入試センター試験では空所に入る語句の選択肢が 4 つ与えられるが、本実験では使用しない。学習データおよびテストデータにおいて空所に補充すべき英単語は 1 語以上 6 語以下である。

\*1 <http://mattmahoney.net/dc/textdata>

### 3.2 性能評価

モデルの性能の評価尺度には BLEU および空所正解率, 空所部分一致率を用いる. ここで空所正解率は式 (1), 空所部分一致率は式 (2) で定義する. モデルの出力のうち “{” と “}” の間に位置する語句を空所に補充する語句とする. モデルの出力に “{” や “}” が出現しない, または 2 回以上出現する場合には不正解と判定する.

空所正解率

$$= \frac{\text{予測した語句と正答が完全一致する問題数}}{\text{全問題数}} \quad (1)$$

空所部分一致率 =

$$\frac{\text{予測した語句が 1 語以上正答に含まれる問題数}}{\text{全問題数}} \quad (2)$$

モデルを 100 回学習したところ, 検証データにおける BLEU が最大となるのは学習回数が 92 回目の時であった. 92 回学習した際のモデルにおける際の検証データおよびテストデータに対する実験結果を表 1 に示す. BLEU に対して空所正解率が低いことから, モデルの出力全体と正答の類似度は高いが, 空所内についての精度は低いことが確認できる.

### 3.3 考察

テストデータに対して, 空所部分の正答および本モデルの予測における出現頻度の上位 5 件とその出現回数を表 2 に示す. 表 2 中の “<unk>” は未知語を表す記号である. 本モデルは予測した単語の確率分布に基づき出力系列を生成するが, その単語はあらかじめ与えられた語彙から選択され, それ以外の英単語は未知語を表す “<unk>” として出力される. 表 2 中の単語の出現回数から確認できるように, 本モデルの予測は一部の単語への偏りが大きい. 空所部分に表れる単語の種類数に関しても正答では 311 種類であるのに対し, 予測では 236 種類であった. また, 検証データに対しても同様の傾向が見られた.

テストデータにおいて空所部分の語句を正解できたのは 324 問中 3 問のみであり, いずれも空所部分は “of” や “in” といった前置詞 1 単語のみの問題であった. その例の一つを表 3 に示す. この例では空所部分は正答できているが, 文全体の完全一致はしていない. 他の 2 問についても同様であり, モデルの出力全体と正解データとの完全一致は 1 問もできなかった. 現在のモデルは Luong らの機械翻訳モデルと同じ構造およびパラメータであるため, 他のモデル構造やハイパー

表 1: モデルの予測精度

	BLEU	空所正解率	空所部分一致率
検証データ	0.6809	0.0740	0.1400
テストデータ	0.7499	0.0093	0.0309

表 2: テストデータにおける出現頻度の上位 5 単語

	正答		予測	
	単語	回数	単語	回数
1	to	14	to	38
2	have	13	of	29
3	been	10	the	28
4	of	9	<unk>	23
5	i	9	in	13

表 3: モデルによる空所部分の予測が正解した例

入力	the fans waited outside the door in the hope { } catching sight of the movie star
予測	the fans waited outside the door in the hope { of } mist sight of the movie star
正答	the fans waited outside the door in the hope { of } catching sight of the movie star

パラメータの検証により, 前後の文脈からより適切に空所に補充する語句を予測できるように改善することが今後の課題である.

## 4 まとめ

本研究では, seq2seq モデルを英文空所補充問題に適用した. text8 コーパスを学習データ, 大学試験センター試験の問題をテストデータとして実験を行ったところ, テストデータにおける BLEU は 0.7499 であったが空所部分を正解できたものは 324 問中 3 問のみであった. 現在のモデルは Luong らの機械翻訳のための seq2seq モデルであるため, 他のモデル構造やハイパーパラメータの検証が今後の課題である.

## 参考文献

- [1] I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. In Neural Information Processing Systems (NIPS). 2014.
- [2] K. Cho, B. V. 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.
- [3] D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473. 2014.
- [4] M. T. Luong, H. Pham, C. D. Manning. Effective Approaches to Attention-based Neural Machine Translation. arXiv preprint arXiv:1508.04025. 2015.
- [5] Z. Xie, A. Avati, N. Arivazhagan, D. Jurafsky, and A. Y. Ng. Neural language correction with character-based attention. arXiv preprint arXiv:1603.09727, 2016.