B-104

転移学習を用いた応答文の口調の変換

我藤 勇樹　謝 孟春　森 徹　村田 充利

1. はじめに

近年、Recurrent Neural Network を用いた sequence to sequence（seq2seq）モデルが機械対話分野で注目を浴びている。seq2seq を用いた対話システムにおいて、応答文の口調は学習に使用したコーパスの影響を強く受けるといわれている。そのため、システムの口調を制御するにはその利用に応じたコーパスを作成する必要があり、膨大なコストがかかることとなる。

転移学習を用いて対話システムのスタイルを制御する研究は、テレビ字幕のコーパスを転移学習用のデータとして seq2seq を訓練する手法がある[1]。本研究では、転移学習を用いてテレビ字幕以外のコーパスで対話システムの口調を変換する手法を検討する。

2. 転移学習による応答文口調の変換

seq2seq とは、ある時系列データを別の時系列データに変換するモデルである。ここでの時系列データとは言語である。このモデルは Encoder-Decoder として 2 つの Long Short Term Memory (LSTM)をもつ (図 1) 。

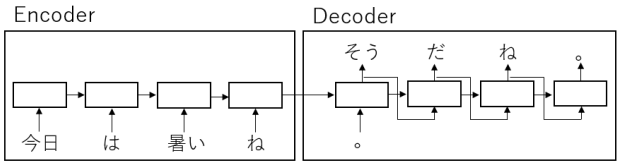

図 1　seq2sseq モデルの一例

転移学習とは、あるドメインで獲得した知識を別のドメインに適応する学習法である。本研究では、転移学習を用いて応答文の口調の変換を行う。そのモデルを図 2 に示す。

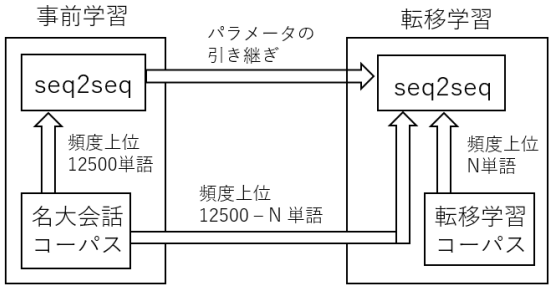

図 2　転移学習による口調の変換モデル

まず、事前学習モデルとして Attention 付き seq2seq を名大会話コーパスで学習させる[2]。この際に過学習防止の為に学習する語彙数に制限を設ける。なお、seq2seq の学習に発話-応答の対を名大会話コーパスで生成する。次に、事前学習によって得たパラメータを引き継ぎ、転移学習コーパスの頻度上位 N 単語と名大会話コーパスの頻度上位 12500 – N 単語を用いて転移学習を行う。

3. 実験と結果

名大会話コーパスで発話-応答対を得るための前処理を行い、約 32,000 個の発話-応答対が得られた。それらを用いて約 12,000 回 seq2seq を事前学習した結果を図 3 に示す。seq2seq の語彙数を 12,500 単語、LSTM のレイヤ数を 2、それらの隠れ層を 256 とした。また、ミニバッチ学習を採用し、バッチサイズを 64 とした。図の横軸は学習回数で、縦軸は perplexity である。perplexity とは、seq2seq における Decoder の予測確率であり、低いほど予測精度が高い。図 3 から perplexity が学習回数約 1500 回まで急速に減少し、その後緩やかに減少していることが分かる。

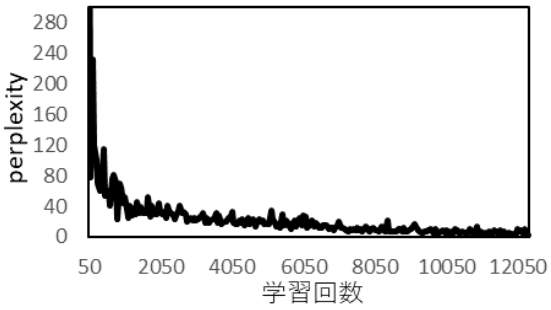

図 3　事前学習の経過

また、実際に会話した例の一部を表 1 に示す。表 1 の A では、短文であるが会話が成立している。それに対して、B では User の発話に脈絡のない応答をしている。

表 1　会話の例

A		B	
	内容		内容
User :	どうも	User :	こんにちは
System :	いえいえどうも	System :	お砂糖じゃない
User :	梅雨入ったね	User :	最近雨多いね
System :	うん	System :	うんいいだよ
User :	雨ばっかりで嫌じゃない？	User :	バイバイ
System :	嫌だね	System :	分かった？

4. 今後の課題

今後の課題として、転移学習用のデータの収集方法が挙げられる。小説のスクレイピングや YouTube 字幕データを用いる予定である。また、会話の不成立を解消するため、データ数の増加や整形方法の見直しも行う必要もある。

参考文献

[1] 赤間, 稲田, 小林, 佐藤, 乾, 転移移学習を用いた対話応答のスタイル制御. 言語処理学会年次大会発表論文集, Vol.23, ROMBUNNNO.B3-3 (2017)

[2] 藤村, 大曽, 大島, 会話コーパスの構築によるコミュニケーション研究, 言語研究の技法：データの収集と分析, pp. 43-72, ひつじ書房 (2011)

和歌山工業高等専門学校