

負の語彙制約に基づくニューラル言い換え生成

梶原智之 (大阪大学データビリティフロンティア機構) kajiwara@jnlp.org

タスク：言い換え生成

本研究では言い換え生成モデルの積極的な編集を改善

入力文と意味的に等価な文を生成するタスク
・テキスト平易化：難解 → 平易の同義文生成
・スタイル変換：カジュアル ↔ フォーマル
テキストにおける**意味以外の情報を制御**できる

機械翻訳と同じ：パラレルコーパス上でseq2seqを訓練
機械翻訳と違う：全ての単語を書き換えるわけではない
タスクごとに基準（難しさ、丁寧さ…）があり、入力文中の基準を満たさない表現を検出し、それらを編集する

一部のみを編集すれば良いというタスクの特徴のために、機械翻訳ベースの言い換え生成モデルはしばしば保守的に振る舞う (= 編集すべき表現まで出力文にコピーしてしまう)

カジュアル → フォーマルの言い換えの例。mama がカジュアルなので言い換えたい。

Source	mama so ugly, she scares buzzards off of a meat wagon.
Reference	Your mother is so unattractive she scared buzzards off of a meat wagon.
ベースライン	mama is so ugly, she scares buzzards off of a meat wagon.
提案手法	The mother is so unattractive that she scares buzzards off of a meat wagon.

提案手法：PMIで言い換え箇所を検出 → その単語を出力しない制約付きビームサーチ

1. 入力文 s_i 中の単語 w のうち、ソーススタイル z との相互情報量が高い語彙 V_i を抽出する

$$V_i = \{w | w \in s_i \wedge \text{PMI}(w, z) \geq \theta\}$$

2. ビームサーチにおいて語彙 V_i を避ける (= 負の語彙制約) ※ 任意の訓練済みモデルに適用可能

評価実験：テキスト平易化（難解→平易）とスタイル変換（カジュアル→フォーマル）

言い換え生成の自動評価。提案手法は2つのモデルと3つのコーパスにおいて、全ての評価尺度を常に改善。

	Newsela					GYAFC-E&M				GYAFC-F&R			
	Add	Keep	Del	BLEU	SARI	Add	Keep	Del	BLEU	Add	Keep	Del	BLEU
RNN-Base	1.8	60.8	22.3	24.1	17.4	31.9	90.0	57.5	71.2	32.9	90.5	61.1	74.7
RNN-PMI	2.8	61.1	36.5	24.7	22.8	33.5	90.0	59.9	71.7	34.3	90.9	63.1	75.9
RNN-Oracle	10.4	82.9	89.9	36.4	40.0	34.8	92.7	72.4	75.2	35.7	93.2	74.6	79.3
SAN-Base	1.8	60.9	23.8	24.0	17.8	34.4	90.0	59.9	71.8	34.5	91.1	63.2	76.7
SAN-PMI	2.5	61.3	38.0	24.6	23.3	35.2	90.0	61.2	72.1	35.3	91.1	64.0	77.0
SAN-Oracle	10.1	82.0	89.4	35.9	39.9	36.6	92.4	71.4	75.1	36.6	92.9	73.7	79.8

Sockeyeでモデルを構築 <https://github.com/aws-labs/sockeye>

- ・RNN: 512次元 × 1層の Bi-LSTM
- ・SAN: 512次元 × 6層の Transformer

Oracle: 検出の上界。参照文に出現しない単語を全て検出。

自動評価 (主) : BLEU, SARI

自動評価 (副) : Add, Keep, Del

→ 適切に追加・保持・削除した単語のF値

各コーパスの文対数。GYAFCのDevとTestはマルチリファレンス (×4)。

	Train	Dev	Test	SOTA (BLEU)
テキスト平易化: Newsela	94,208	1,129	1,077	24.3
スタイル変換: GYAFC-E&M	52,595	2,877	1,416	71.4
スタイル変換: GYAFC-F&R	51,967	2,788	1,332	74.5

あまり多くの制約をかけるべきではない

