

訂正文の流暢性向上を目的とした 系列タグ付け文法誤り訂正器の強化学習手法



五藤巧¹, 渡辺太郎¹ ¹ 奈良先端科学技術大学院大学

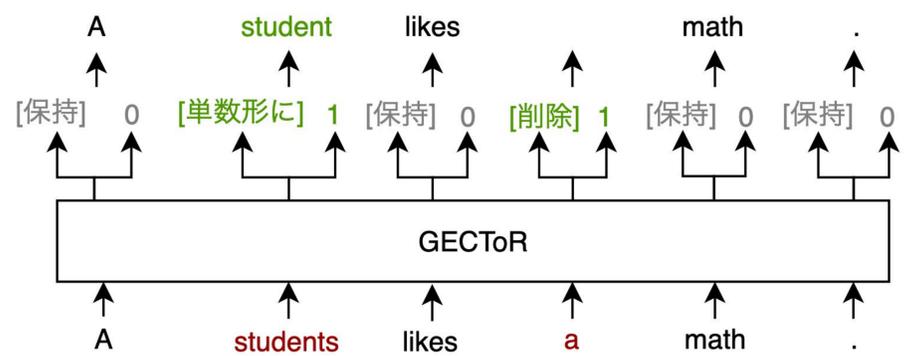
系列タグ付けによる文法誤り訂正

● GECToR [Omelianchuk+ 20]

- 言語情報を考慮したタグを推定
- タグ推定と誤り検出のマルチタスク学習

● 利点

- 説明性: タグ→フィードバック, 誤り検出→確信度
- 高速な推論: 非自己回帰モデル



Omelianchuk et al, GECToR – Grammatical Error Correction: Tag, Not Rewrite, BEA2020

GECToRの流暢性

● GECToRの流暢性は系列変換モデルと比べて低い

● 評価尺度

- GLEU: n-gramに基づく参照あり尺度
- SOME: 人手評価結果に直接最適化した参照なし尺度
- GPT-2 PPL: GPT-2のperplexity

流暢な訂正のベンチマーク: JFLEG-dev

モデル	↑GLEU	↑SOME流暢性	↓GPT-2 PPL
Kiyonoら(系列変換)	56.2	0.794	278.9
Kanekoら(系列変換)	56.1	0.794	213.9
GECToR(系列タグ付け)	54.3	0.787	308.7

流暢性の向上を目的とした強化学習

● なぜ流暢性が低いのか?

→最適化に訂正文が考慮されていない

● 訂正文の流暢性を報酬として強化学習

- 貪欲な推定系列に対する損失を新たに考慮
- 流暢性の観点から報酬を与えて最適化

● 報酬: 負のGPT-2のperplexity

- -300から-1の範囲でmin-max正規化
- ベースライン報酬: ミニバッチ内の平均報酬

● 従来の損失関数:

$$\mathcal{L}_{mle} = - \sum_{i=1}^{|\mathbf{x}|} (\log P^t(y_i^t | x_i, \mathbf{x}) + \log P^d(y_i^d | x_i, \mathbf{x}))$$

正解ラベルに対する損失

● 強化学習による損失関数:

$$\mathcal{L}_{rl} = - (r(\hat{t}) - \hat{r}) \sum_{i=1}^{|\mathbf{x}|} (\log P^t(\hat{y}_i^t | x_i, \mathbf{x}) + \log P^d(\hat{y}_i^d | x_i, \mathbf{x}))$$

流暢性の観点からの報酬 貪欲な推定系列に対する損失

● 最終的な損失関数:

$$\mathcal{L} = \lambda \mathcal{L}_{mle} + (1 - \lambda) \mathcal{L}_{rl}$$

訂正性能の向上 流暢性の向上

実験

- 3段階の訓練: 1段階目は擬似誤り(8M文対, 損失 \mathcal{L}_{mle}) / 2・3段階目は実誤り(560K文対, 損失 \mathcal{L})

● 結果: 訂正文の流暢性の向上を確認

● $\lambda=1$ と $\lambda=0.5$ との比較

- 提案法が流暢性の向上に寄与
- 訂正性能はほぼ維持

● $\lambda=0.5$ と系列変換モデルとの比較

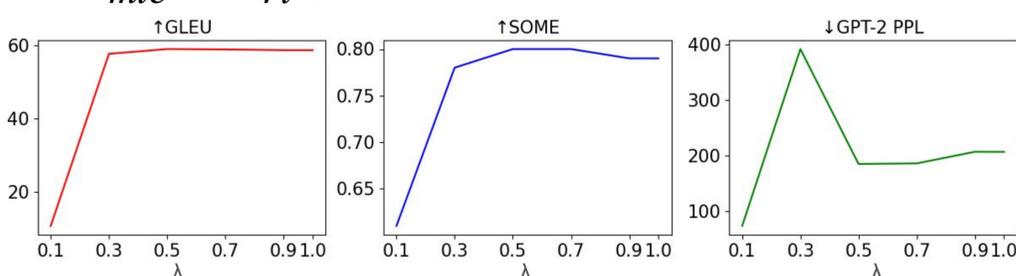
- 流暢性は競合的に
- 訂正性能は提案手法が劣る

	JFLEG-test			CoNLL-2014
	↑GLEU	↑SOME流暢性	↓GPT-2 PPL	F0.5
Kiyonoら	61.0	0.797	193.1	62.0
Kanekoら	61.0	0.797	175.5	62.7
GECToR ($\lambda=1$)	58.6	0.799	207.3	61.4
提案法 ($\lambda=0.5$)	58.9	0.805	185.5	60.8

分析

● λ の影響: $\lambda=0.5$ 付近で高い性能を示す

- \mathcal{L}_{mle} と \mathcal{L}_{rl} を概ね同じ重みで考慮すべき



※ $\lambda=0.1$ の場合のGLEUについて、論文の表5では58.5と記載していますが、正しくは10.8です。訂正してお詫びいたします。

● 定性分析: 内容語の語彙選択の訂正が増加

- 一方, 意味保存性に欠けた訂正も
 - 報酬関数が意味保存性の観点を含まないから?

	入力文	GPT-2のperplexity
	... gives me many new abilities in <u>filmmaking</u> .	269.8
$\lambda=1$... gives me many new abilities in <u>making</u> .	175.4
$\lambda=0.5$... gives me many new abilities in <u>photography</u> .	135.3