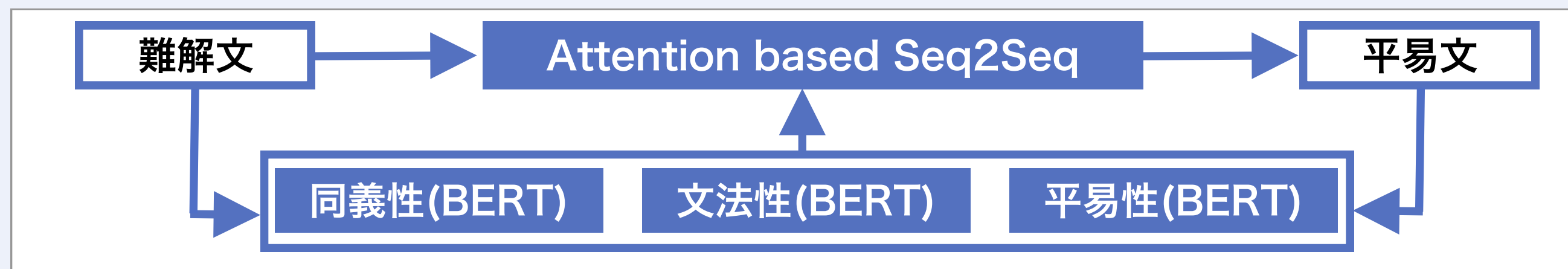


# 深層強化学習によるテキスト平易化における文難易度報酬の適用

中町 礼文, 梶原 智之, 荒瀬 由紀 (大阪大学)

## 1. 概要

- テキスト平易化: 難解な文を, 平易な文に書き換えるタスク
- 文の難易度を直接評価する訓練手法を本論文で初めて提案
- 文の難易度を考慮した強化学習で, 平易化モデルの訓練を行った



## 2. 文生成の強化学習

Attention based Seq2Seqに対して, REINFORCEを適用.  
REINFORCE: 生成モデルのCrossEntropyを即時報酬  $R_t$  で重み付けする.

$$\mathcal{L}_{RE} = - \sum_{k=1}^K \sum_{i=1}^M R_t y_i^k \log p_{dec}(y_i^k | y_{1:i-1}^k, \mathbf{h}^k)$$

$$R_t = r(\cdot) - b_t$$

$r(\cdot)$  は, 出力文から計算される報酬の総和. 文生成の強化学習では,  $r(\cdot)$  を課題に応じて設計. テキスト平易化では一般的に, 入出力文の**同義性**と出力文の**文法性**と**平易性**が用いられる.

## 3. 報酬設計

	提案手法	既存手法 (DRESS)
同義性	STS-Bの文対から意味類似性を推定	文ベクトルのコサイン類似度
文法性	GUGから文法的正しさを推定	Newselaの平易文のLSTM言語モデル
平易性	Newselaの文から学年ラベルを推定	SARI (入出力間の語句の追加・保持・削除のF値)

提案手法では**文の難易度を考慮できる**.  
提案手法では, BERTのファインチューニングでそれぞれの文評価器を作成した. 報酬の総和として, 文評価器の出力を正規化した値の平均値を用いた.

## 4. 実験結果

- 提案手法は, 平易性の評価値が最も高い
- 提案手法は, 同義性, 文法性, 平易性の3つをバランスよく高めている
- 報酬推定器は概ね正しく予測できている

	Train	Dev	Test	Label	Pearson r
同義性(STS-B)	5,749	1,500	1,379	[0, 5]	0.897
文法性(GUG)	1,518	747	754	[1, 4]	0.655
平易性(Newsela)	94,208	1,129	1,077	[2, 12]	0.762

	BLEU ↑	SARI ↑	FKGL ↓	同義性 ↑	文法性 ↑	平易性 ↑	平均報酬 ↑
DRESS	<b>23.21</b>	<b>27.37</b>	<b>4.13</b>	0.732	<b>0.952</b>	0.606	0.763
EncDecA	22.40	25.61	5.50	0.734	0.916	0.617	0.756
BERT-RL	23.02	25.31	4.86	<b>0.785</b>	0.881	<b>0.641</b>	<b>0.769</b>

入力(難解文)	They are tired and it shows in their voices, but they're still on the freedom highway.
参照(平易文)	Their voices sound tired.
EncDecA	They are tired and it shows in their voices, but they're still on the freedom.
BERT-RL	They are tired and it shows in their voices.