Text Simplification with Reinforcement Learning using Supervised Rewards on Grammaticality, Meaning Preservation, and Simplicity

Akifumi Nakamachi, Tomoyuki Kajiwara, Yuki Arase
Osaka University

Background
Text Simplification

Text simplification
- The text-to-text generation tasks that rewrites complex sentence simpler
- Applications: language learning support, pre-processing for other NLP tasks

The ideal text simplification:
Rewrite complex sentence in a grammatically correct and simple manner while preserving the meaning.

Input		Output		
This is an inexpensive pen.	→	Pen.	✕	cannot keep meaning
	→	This is an inexpensive pen.	✕	No rewrite
	→	This is a cheap pen.	○	

Background
Problem in Neural Text Simplification Approach

Exposure bias:
The model is not exposed to its own errors during training

Loss-evaluation mismatch:
Training objective (loss) is different from evaluation at an inference

Previous Studies

Text Simplification with Reinforcement Learning

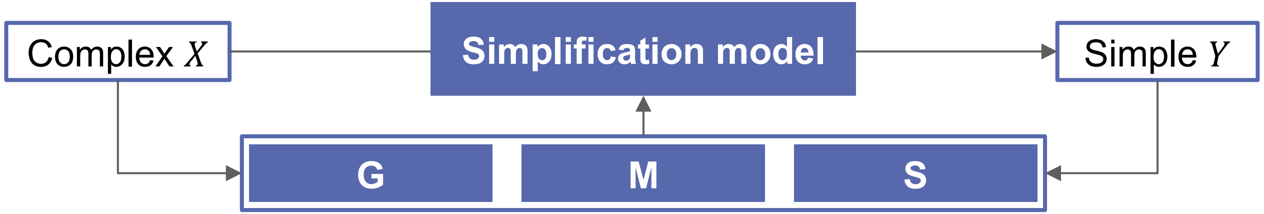
1. Pre-training the model with difficult sentence X and simple sentence Y

$$\mathcal{L}_{\mathrm{C}} = -\sum_{t=1}^{|Y|} \log P(y_{t+1}|y_{1\cdots t}, X)$$

2. Reinforcement learning to increase the **evaluation** R **of the output sentence**

$$\mathcal{L}_{\mathrm{R}} = -\sum_{t=1}^{|Y|} r(h_t) \log P(y_{t+1}|y_{1\cdots t}, X) \qquad r(h_t) = R(\cdot) - b(h_t)$$

The R is the sum of the evaluations of **G**rammaticality, **M**eaning preservation, and **S**implicity calculated from X and Y



[1] Zhang and Lapata. Sentence Simplification with Deep Reinforcement Learning. EMNLP 2017.
[2] Zhao et al. Semi-Supervised Text Simplification with Back-Translation and Asymmetric Denoising Autoencoders. AAAI 2020.

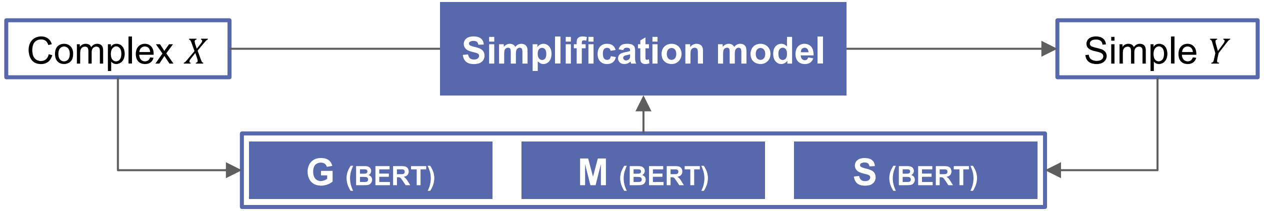
Proposed Method

BERT-based Rewards for Accurate Estimation of Human Sense

Create the estimators for the three viewpoints of simplification
by supervised learning using BERT [3]

- **G**rammaticality: GUG [4] (grammar evaluation data set)
- **M**eaning perservation:STS-B [5] (synonymy evaluation data set)
- **S**implicity: Newsela [6] (simplification data set)

By training with rewards that are optimized for human evaluation,
simplification by the system approaches human senses



[3] Devlin et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL 2019.
[4] Heilman et al. Predicting Grammaticality on an Ordinal Scale. ACL 2014.
[5] Wang et al. GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding.EMNLP 2018.
[6] Xu et al. Problems in Current Text Simplification Research: New Data Can Help. TACL 2015.

Comparison with Existing Methods

Methods	G	M	S
Zhang and Lapata [1]	Sentence generation probability of Y by language model	Cosine similarity between sentence vector of X, Y	SARI [7,8] (simplification metric)
Zhao et al [2]	Sentence generation probability of Y by language model	Cosine similarity Between averaged word vector of X, Y	FKGL [9] (readability metric)
Ours	BERT based estimator		

[7] Xu et al. Optimizing Statistical Machine Translation for Text Simplification. TACL 2016
[8] Alva-Manchego et al. ASSET: A Dataset for Tuning and Evaluation of Sentence Simplification Models with Multiple Rewriting Transformations. ACL 2020.
[9] Kincaid et al. Derivation Of New Readability Formulas (Automated Readability Index, Fog Count And Flesch Reading Ease Formula) For Navy Enlisted Personnel. DTIC Document 1975.

Experiments

A comparative experiment of the sentence evaluators

The correlation coefficients between each estimator and the human evaluations of each method are summarized

Compared with the other estimators, the proposed method has **higher correlation with the human evaluations**

Method	G	M	S
Zhang and Lapata.	0.041	-0.135	0.034
Zhao et al.	0.379	-0.135	0.175
Ours	**0.726**	**0.846**	**0.473**

Result of the experiment

	Train	Dev	Test	Label
G (GUG)	1,518	747	754	1~4
M (STS-B)	5,749	1,500	1,379	0~5
S (Newsela)	94,208	1,129	1,077	2~12

Data sets information

Experiment

End-to-End Evaluation on Text Simplification

- Data set: Newsela (Simplification parallel data set)
- Simplification model: EncDecA (encoder-decoder model with attention)
- Automatic evaluations: Simplification metrics (BLEU, SARI, FKGL) and averaged estimation scores of BERT based estimators
- Human evaluations:

  The average rankings of the output sentences of the 3 systems with reinforcement learning,

  Ranked manually in the order of their simplicity while retaining appropriate meaning

Experimental Results

End-to-End Evaluation on Text Simplification

The reward was the highest.

In the human evaluation, our proposed method rated as the simplest while preserving meaning enough of the three reinforcement-based text ($p < 0.05$)

	BLEU	SARI	FKGL	G	M	S	Reward	Rank by Human
EncDecA	21.57	37.64	4.39	0.862	**0.681**	0.648	0.730	n/a
Zhang and Lapata.	23.30	**39.24**	**4.09**	**0.878**	0.659	**0.663**	0.734	1.69
Zhao et al.	**23.42**	39.20	4.24	**0.878**	0.662	0.662	0.734	1.91
Ours	23.14	28.70	4.53	**0.878**	0.678	0.653	**0.736**	**1.45***

Experimental Results
Example of Generated Text

Complex	They are tired and it shows in their voices , but they ’re still on the freedom highway.
Simple (Reference)	Their voice sound tired.
EncDecA	They are tired and it shows in their voices , but they ’re still on the freedom.
Zhang and Lapata	They are tired.
Zhao et al.	They are tired.
Ours	They are tired and it shows in their voices.

Conclusion

- We proposed BERT based supervised estimators for text simplification
- Comparative experiments of estimators showed that our estimators have more correlate on with human senses than existing methods
- We applied our proposed estimators as a reward for reinforcement learning
- Comparative experiment of reinforcement text simplification, the proposed method is the highest in averaged overall reward and is significantly different in terms of human evaluation.