Readability-guided Idiom-aware Sentence Simplification (RISS) for Chinese

Jingshen Zhang, Xinglu Chen, Xinying Qiu*, Zhimin Wang, Wenhe Feng

Guangdong University of Foreign Studies, Guangzhou, China

Abstract

- **Definition**

Sentence simplification (SS) aims to reduce the linguistic complexity of a sentence while preserving its meaning.

- **Challenge**

However, Chinese sentence simplification faces two main challenges: *the scarcity of parallel training data and the prevalence of idioms.*

- **Proposal**

We propose *Readability-guided Idiom-aware* Sentence Simplification (RISS), a novel framework combining data augmentation tech.

Methodology I

RPS: Readability-guided Paraphrase Selection

Utilizing *pairwise readability differences and Levenshtein/syntactic/semantic similarities* to automatically mine high-quality training data.

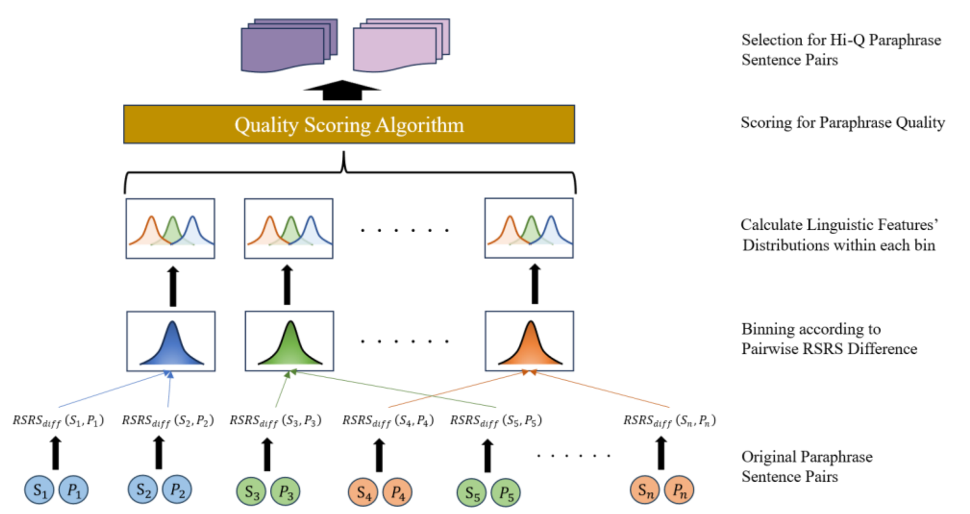

Figure 1: The architecture of the proposed RPS framework

Methodology II

IAS: Idiom-aware Simplification

Introducing *an idiom-specific loss component* to encourage accurate representations and simplification of idiomatic phrase.

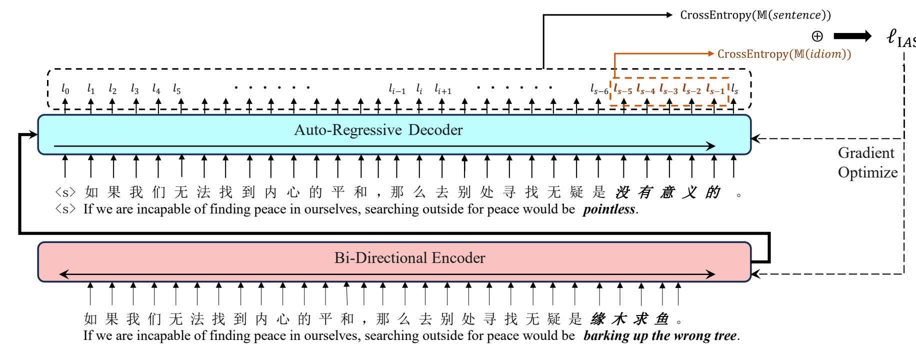

Figure 2: The architecture of the proposed IAS framework

Combined Strategy Framework

RISS

Using *multi-stage or multi-task* to combine RPS and IAS .

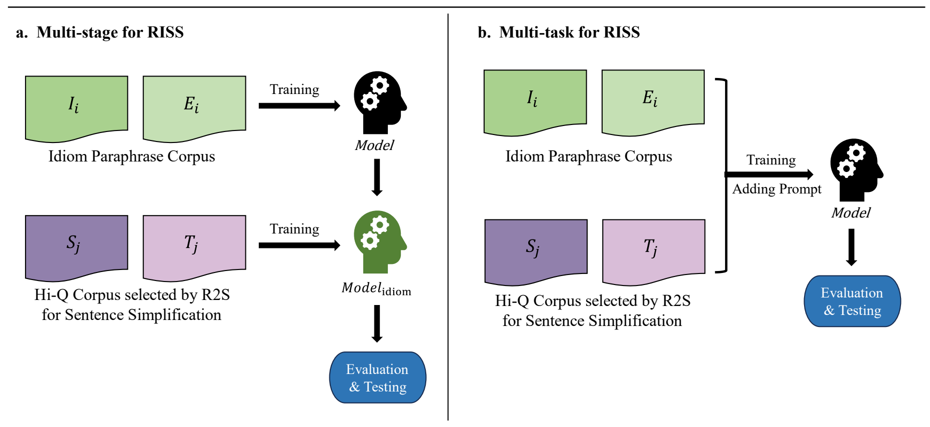

Figure 3: The RISS module architecture with multi-stage (left) and multi-task (right) strategies, where I_i , E_i denote idioms and their corresponding explanations respectively, and S_j , T_j denote source and target sentences selected by RPS respectively.

Main Result & Ablation Study

Method	CSS $Sari_{char}$	CSS $Sari_{word}$	CSS $B.S.$	CSS $BLEU$	MCTS $Sari_{char}$	MCTS $Sari_{word}$	MCTS $B.S.$	MCTS $BLEU$
Identity	29.08	27.61	0.88	88.77	25.06	22.37	0.90	84.75
Truncation	32.95	33.18	0.79	76.36	25.21	21.85	0.87	61.6
Gold Reference	46.72	45.71	0.96	65.31	50.12	48.11	0.90	61.62
Unsupervised (Zero-Shot) Method								
Lu et al. (2021)	36.27	33.39	-	63.47	-	40.37	-	48.72
Translate Training (2023)	36.02	34.44	-	71.41	-	28.30	-	**82.20**
Cross-Lingual Pseudo (2023)	-	-	-	-	-	38.49	-	63.06
Gpt-3.5-turbo (2023)	31.95	28.92	0.83	42.22	-	42.39	-	49.22
$\text{RISS}_{\text{multi-stage}}$	40.95	39.36	0.81	81.92	44.37	42.71	**0.90**	79.94
$\text{RISS}_{\text{multi-task}}$	**41.68**	**40.52**	**0.87**	**84.05**	**46.23**	**44.36**	**0.90**	80.83
Supervised (Few-Shot) Method								
mT5-large (2023)	37.57	35.97	-	74.71	-	-	-	-
Gpt-3.5-turbo (2023)	39.32	36.57	0.85	60.57	-	-	-	-
$\text{RISS}_{\text{multi-stage+labeled data}}$	**44**	**43.1**	**0.88**	88.48	47.41	45.94	**0.90**	79.94
$\text{RISS}_{\text{multi-task+labeled data}}$	43	42.35	**0.88**	**88.59**	**48.49**	**46.42**	**0.90**	**81.12**

Table 2: Main results of the experiment. For a comprehensive analysis, we calculate the missing metrics for publicly available system-generated results provided by previous studies; otherwise, we use "-" to fill in the missing values. **Bold** indicates the best result, and underline indicates the second-best result. For detailed experiment parameters, please see Appendix B.

Method	CSS $Sari_{char}$	CSS $Sari_{word}$	CSS $B.S.$	CSS $BLEU$	MCTS $Sari_{char}$	MCTS $Sari_{word}$	MCTS $B.S.$	MCTS $BLEU$
Unsupervised (Zero-Shot) Method								
Raw Paraphrase	36.45	35.22	0.806	86.68	41.27	39.8	**0.908**	81.1
RPS	40.54	38.98	0.867	82.51	43.89	42.29	0.90	78.24
$\text{IAS}_{\text{without-loss}}$	36.23	35.39	0.874	89.17	45.84	43.72	0.90	**81.73**
$\text{IAS}_{\text{with-loss}}$	36.56	35.89	**0.878**	**89.66**	**46.31**	**44.76**	0.901	81.18
$\text{RISS}_{\text{multi-stage}}$	40.95	39.36	0.811	81.92	44.37	42.71	0.899	79.94
$\text{RISS}_{\text{multi-task}}$	**41.68**	**40.52**	0.873	84.05	46.23	44.36	0.899	80.83
Supervised (Few-Shot) Method								
Bart	40.21	39.39	0.886	88.13	44.8	43.43	0.899	81.48
Raw Paraphrase	41.32	40.33	0.868	**89.45**	45.13	44.28	**0.905**	81.49
RPS	43.76	42.68	0.87	87.84	46.47	45.48	0.903	80.94
$\text{IAS}_{\text{without-loss}}$	41.51	40.73	0.877	88.34	47.37	45.71	0.903	**81.68**
$\text{IAS}_{\text{with-loss}}$	42.99	42.26	**0.902**	88.63	47.62	46.11	0.895	81.15
$\text{RISS}_{\text{multi-stage}}$	**44**	**43.1**	0.875	88.48	47.42	45.94	0.903	81.57
$\text{RISS}_{\text{multi-task}}$	43	42.35	0.876	88.59	**48.49**	**46.42**	0.904	81.12

Table 3: Ablation Study on CSS and MCTS. **Bold** indicates the best result, and underline indicates the second-best result.

Conclusions

- ***Addressing two major challenges*** in Chinese sentence simplification.
- ***Advancing the SOTA performance*** on two Chinese sentence simplification datasets.

audbut0702@163.com xy.qiu@foxmail.com