

---

# Edit Distance Based Curriculum Learning for Paraphrase Generation

Sora Kadotani\* Tomoyuki Kajiwara\*\* Yuki Arase\* Makoto Onizuka\*

\*Graduate School of Information Science and Technology, Osaka University

\*\*Graduate School of Science and Engineering, Ehime University

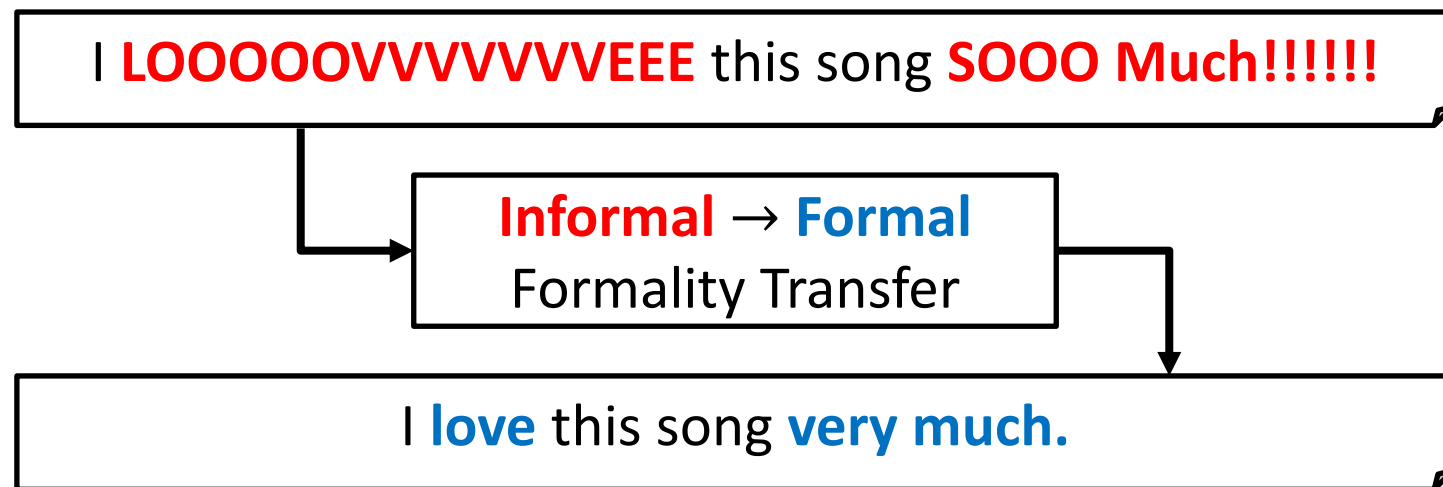


OSAKA UNIVERSITY

# Paraphrase Generation

---

- Can be formulated as a monolingual machine translation problem
- Application examples
  - English education support
  - Preprocessing for machine translation



# Curriculum Learning

---

- Training a model starting from easy samples and gradually moving to difficult ones
- Curriculum learning for neural machine translation [1, 2]
  - Improve translation quality
  - The metric of difficulty: Sentence length, Word rarity

**Easy**

Thank you.

**Medium**

Thank you very much.

**Difficult**

Thank you for your  
helping me with my work.



**Training Time**

[1] Platanios et al. (NAACL 19) Competence-based Curriculum Learning for Neural Machine Translation

[2] Liu et al. (ACL 20) Norm-Based Curriculum Learning for Neural Machine Translation

# Difficulty of Paraphrase Generation

---

- Paraphrases require a few transformations:  
Copy almost all the input sentence's words (**Easy**)
- Paraphrases require drastic transformations:  
Require complex rewriting operations (**Difficult**)
- We propose to estimate the difficulty of transformations by **edit distance** between sentence pair

---

Source Sentence

Target Sentence

---

Their first two albums were **pretty** good.      Their first two albums were **very** good.

no where there is no such thing

That does not exist.

---

# Edit Distance

---

- Number of edit operations required to convert the word sequence
- Small edit distance: **Easy**
- Large edit distance: **Difficult**

**Edit distance = 1**

---

Source Sentence

Target Sentence

---

Their first two albums were **pretty** good. Their first two albums were **very** good.

---

no where there is no such thing

That does not existst.

---

**Edit distance = 7**

# Proposed Method

---

## Edit Distance Based Curriculum Learning

- Apply curriculum learning to paraphrase generation
- Use the existing curriculum learning framework [1]
- The metric of difficulty: Edit distance

# Experiment: Setup

---

## Evaluate the quality of paraphrase generation

- Dataset: GYAFC [3]

- Model: Transformer

- Evaluation metric: BLEU

- Comparison methods

- Baseline: Without curriculum learning
- CL-SL: Curriculum learning with sentence length
- SL-WR: Curriculum learning with word rarity
- CL-ED: Curriculum learning with edit distance

	Train	Dev	Test
E&M	209,124	2,877	1,416
F&R	209,124	2,788	1,332

[3] Rao and Tetreault (NAACL 18) Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer

# Experiment: Results

---

## Evaluate the quality of paraphrase generation

- Only CL-ED outperformed Baseline on both domains
- The proposed method was effective

	Informal → Formal	
	E&M	F&R
Source	49.19	50.94
Baseline	69.81	75.02
CL-SL	69.83	74.90
CL-WR	70.05	74.62
CL-ED	<b>70.34</b>	<b>75.41</b>



# Analysis: Setup

---

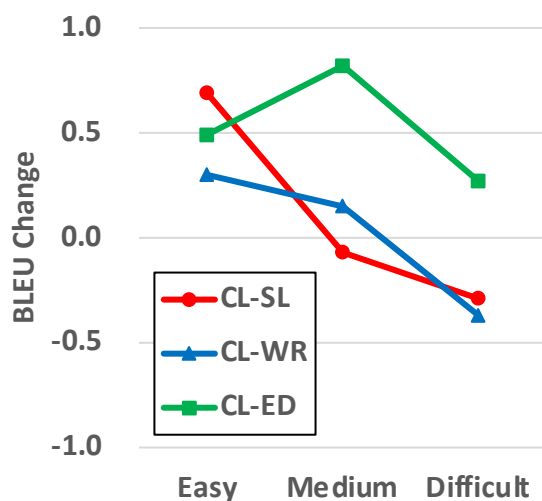
## Investigate which types of sentences are improved

- Procedures
  - Divide the test set based on difficulty levels
  - Compute a BLEU score of each class
  - Calculate improvements over Baseline
- Comparison methods
  - CL-SL: Curriculum learning with sentence length
  - SL-WR: Curriculum learning with word rarity
  - CL-ED: Curriculum learning with edit distance

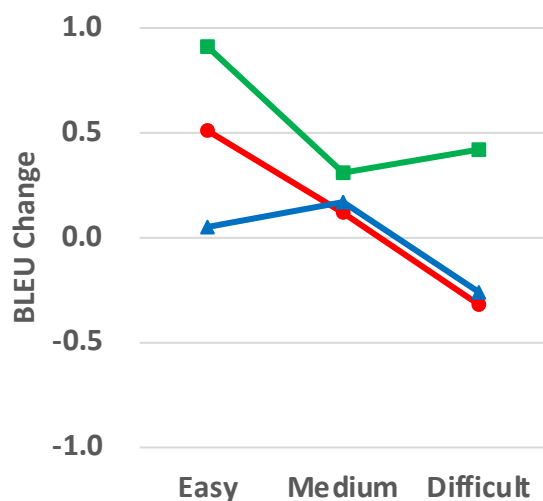
# Analysis: Results

## Investigate which types of sentences are improved

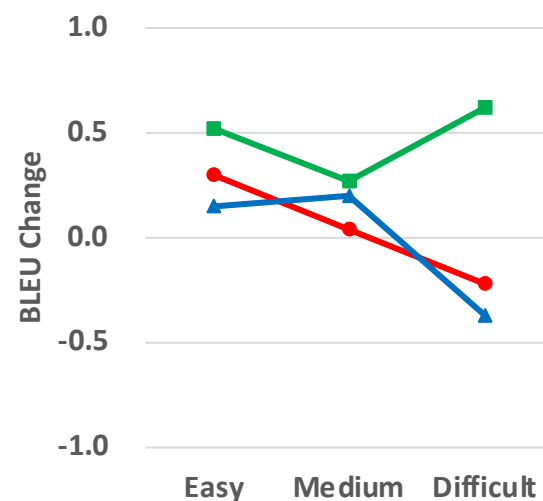
- CL-SL and CL-WR degraded the quality of difficult samples
- CL-ED improved the quality of all classes regardless of difficulty levels



(a) Sentence Length



(b) Word Rarity



(c) Edit Distance

# Conclusion

---

- Summary
  - Propose the **edit distance** as a difficulty metric in curriculum learning
  - The proposed method is effective in paraphrase generation
  - The proposed method improves the quality regardless of difficulty levels
- Future work
  - Curriculum learning can be applied to any task
    - Apply curriculum learning to transfer learning