Dictionary-Guided Editing Networks for Paraphrase Generation

Shaohan Huang† , Yu Wu‡ , Furu Wei† , Ming Zhou†

†Microsoft Research, Beijing, China

‡State Key Lab of Software Development Environment, Beihang University, Beijing, China

AAAI 2018

Background

- Paraphrase generation aims to generate restatements of the meaning of a text or passage using other words
- Application - information retrieval, question answering, and conversation systems
- Prakash et al. (2016) employ the residual recurrent neural networks for paraphrase generation, that is one of the first major words that uses a deep learning model for this task
- Gupta et al. (2017) propose a combination of VAE and sequence-to-sequence model to generate paraphrase

Introduction

- Existing work on paraphrase generation focuses on generating paraphrase sentences from scratch (e.g. seq2seq *[Lin+, 2014]*)

- However, an intuitive way for a human to write paraphrase sentences by replacing words or phrases in the original sentence

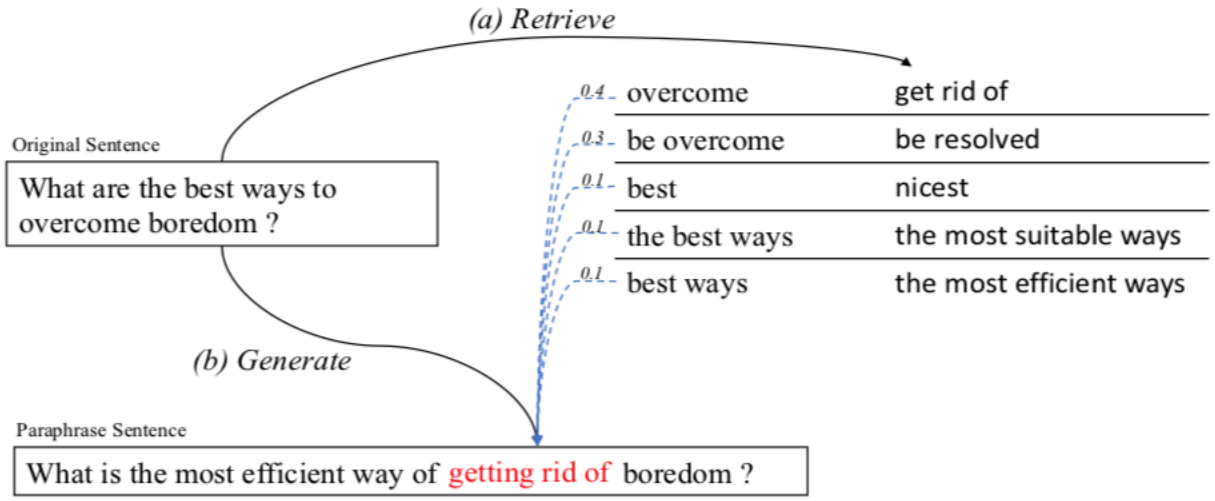
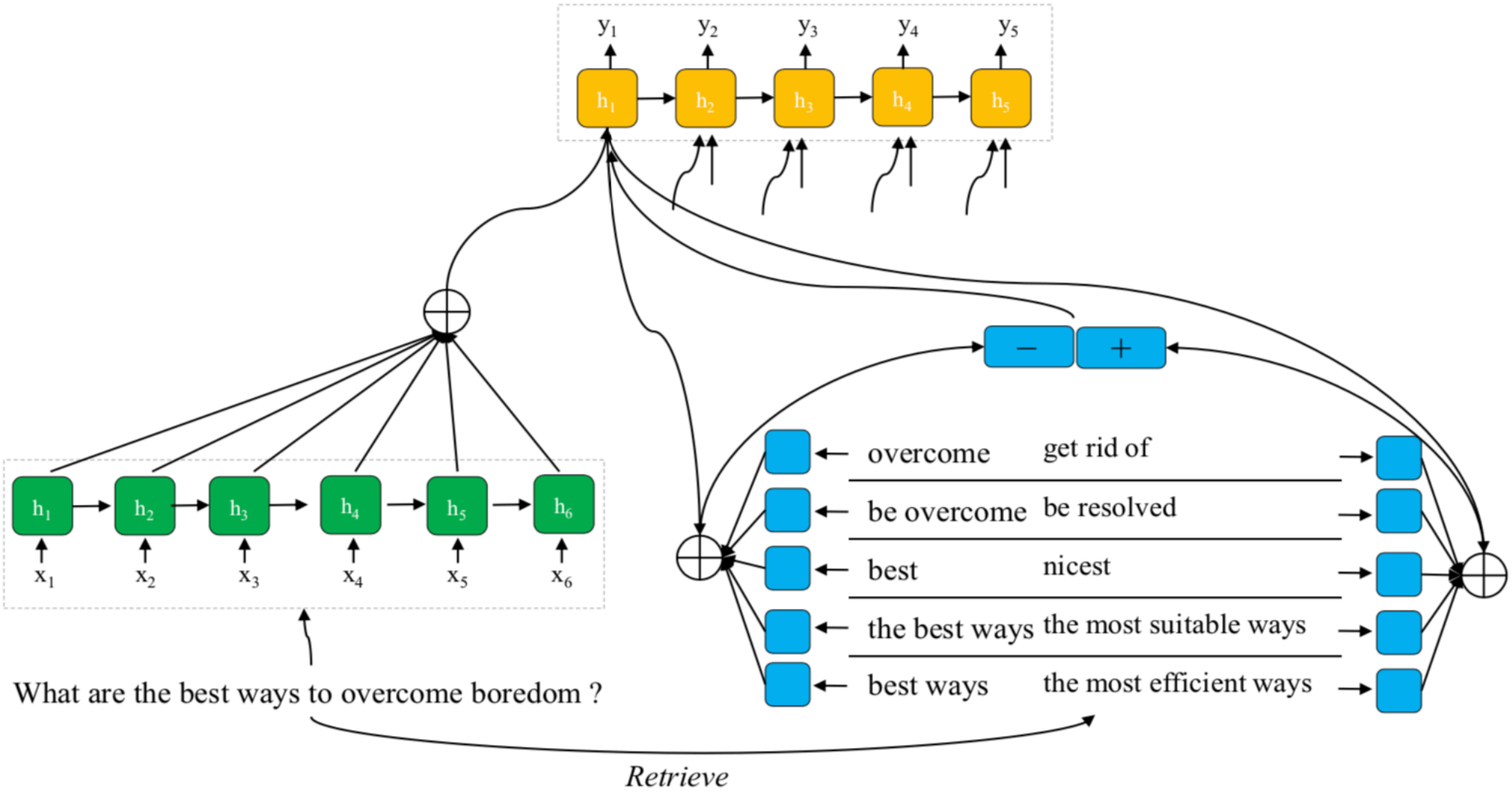


Figure 1: The dictionary-guided editing networks model first retrieves a group of paraphrased pairs and then generates a paraphrase using the original sentence as a prototype.

- Novelty: it jointly learns the (i) selection of the appropriate paraphrase pairs and (ii) generation of fluent sentences.

Methodology



Methodology – Retrieve

- Original Sentence X “What are the best ways to overcome boredom ? ”
- Paraphrased pairs D such as (“overcome”, “get rid of”), (“the best ways”, “the most suitable ways”), and (“the best ways”, “the most efficient ways”)

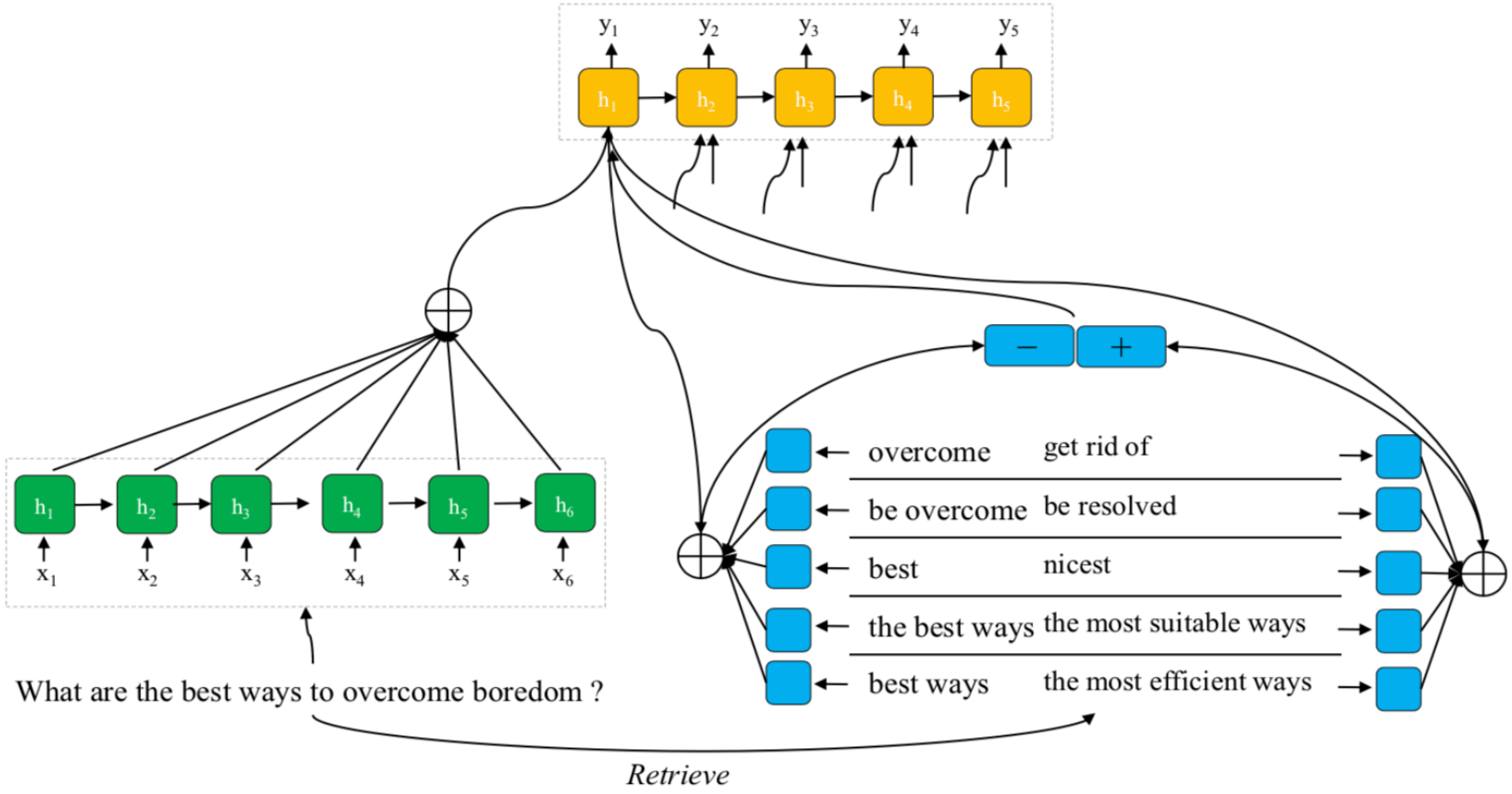
```
132882  [VBD] ||| combined ||| blends ||| PPDB2.0Score=3.55781 PPDB1.0Score=15.479350 -logp(
        LHS|e1)=3.39490 -logp(LHS|e2)=3.48857 -logp(e1|LHS)=9.97208 -logp(e1|e2)=5.68093 -logp(
        e1|e2,LHS)=1.11850 -logp(e2|LHS)=14.18324 -logp(e2|e1)=9.79842 -logp(e2|e1,LHS)=5.32966
        AGigaSim=0.51410 Abstract=0 Adjacent=0 CharCountDiff=-2 CharLogCR=-0.28768 ContainsX=0
        Equivalence=0.269878 Exclusion=0.001897 GlueRule=0 GoogleNgramSim=0.03067 Identity=0
        Independent=0.230379 Lex(e1|e2)=62.39060 Lex(e2|e1)=62.39060 Lexical=1 LogCount=0.69315
        MVLSASim=NA Monotonic=1 OtherRelated=0.183808 PhrasePenalty=1 RarityPenalty=0.00002
        ReverseEntailment=0.314039 SourceTerminalsButNoTarget=0 SourceWords=1 TargetComplexity=0.99275
        TargetFormality=0.99686 TargetTerminalsButNoSource=0 TargetWords=1 UnalignedSource=0
        UnalignedTarget=0 WordCountDiff=0 WordLenDiff=-2.00000 WordLogCR=0 ||| 0-0 ||| ReverseEntailment
132883  [VBD] ||| overcome ||| addressed ||| PPDB2.0Score=3.55781 PPDB1.0Score=16.528730 -logp(
        LHS|e1)=9.65618 -logp(LHS|e2)=1.72090 -logp(e1|LHS)=16.22013 -logp(e1|e2)=8.92465 -logp(
        e1|e2,LHS)=11.91261 -logp(e2|LHS)=6.96427 -logp(e2|e1)=7.60408 -logp(e2|e1,LHS)=2.65676
        AGigaSim=0.59570 Abstract=0 Adjacent=0 CharCountDiff=1 CharLogCR=0.11778 ContainsX=0
        Equivalence=0.012291 Exclusion=0.000686 GlueRule=0 GoogleNgramSim=0.04294 Identity=0
        Independent=0.418977 Lex(e1|e2)=62.90141 Lex(e2|e1)=62.90141 Lexical=1 LogCount=0 MVLSASim=NA
        Monotonic=1 OtherRelated=0.321855 PhrasePenalty=1 RarityPenalty=0.36788 ForwardEntailment=0.246191
        SourceTerminalsButNoTarget=0 SourceWords=1 TargetComplexity=0.99429 TargetFormality=0.99876
        TargetTerminalsButNoSource=0 TargetWords=1 UnalignedSource=0 UnalignedTarget=0 WordCountDiff=0
        WordLenDiff=1.00000 WordLogCR=0 ||| 0-0 ||| Independent
132884  [VBN] ||| redistributed ||| split ||| PPDB2.0Score=3.55781 PPDB1.0Score=19.245440 -logp(
        LHS|e1)=0.18102 -logp(LHS|e2)=1.50019 -logp(e1|LHS)=11.52214 -logp(e1|e2)=11.09932 -logp(
```

Ranking Function

$$score_r = \sum_{w \in o_i \cap x} tf_w \cdot idf_w + score'(o_i, p_i)$$

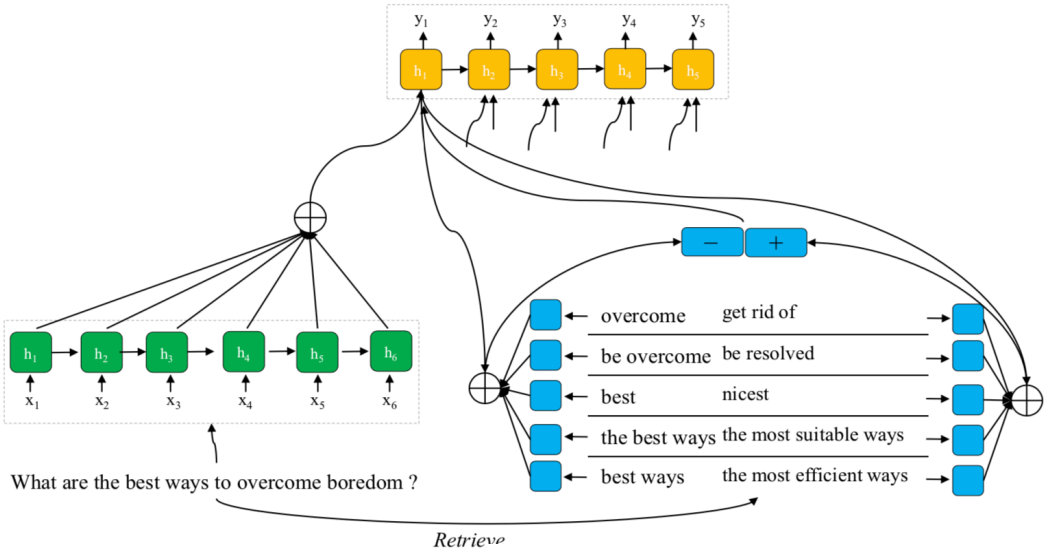
- 取10*M 个cadidates rank最后取top M 个

Methodology – Dictionary-Guided Editing



For each paraphrased pair in $\mathcal{E}$, we employ the same encoding method and obtain $2 \times M$ vectors $\mathcal{E}' = \{(o_r^i, p_r^i)\}_{i=1}^{M}$. In the next section, we will introduce leveraging our paraphrased dictionary to generate a paraphrase.

Methodology – Dictionary-Guided Editing



$$\tilde{\mathbf{h}}_\mathbf{t} = tanh(\mathbf{W_c} \cdot (\mathbf{h_t} \oplus \mathbf{c_t} \oplus \mathbf{c'_t}))$$

- C'_t is computed as the weighted average of the original hidden states

the weighted average of o_r^i and p_r^i as follows:

$$\mathbf{c_t} = \sum_{i=1}^{M} \mathbf{a_{t,i}} \cdot o_r^i \oplus \sum_{i=1}^{M} \mathbf{a'_{t,i}} \cdot p_r^i$$

$$\mathbf{y_t} = softmax(\mathbf{W_y}[\mathbf{y_{t-1}} \oplus \tilde{\mathbf{h}}_\mathbf{t} \oplus \mathbf{c_t} \oplus \mathbf{c'_t}] + \mathbf{b_y})$$

Experiments

- Objective function minimizing the negative log-likelihood

$$\mathcal{J} = -\log(p(\mathbf{y}^*|\mathbf{x}, \mathcal{E}'))$$

- **Dataset**
  - MSCOCO (Lin et al. 2014) - a large-scale captioning dataset which contains **five human annotated captions** of over 120K images
  - **Quora** dataset is related to the problem of identifying duplicate questions. It consists of over 400K potential question duplicate pairs.
- **Evalutions**: BLEU (Papineni et al. 2002) and METEOR (Lavie and Agarwal 2007).
  (Correlate human judgments in paraphrase generation, Wubben+, 2010; Madnani+, 2012)

Results

Table 1: Results on MSCOCO dataset. Higher BLEU and METEOR score is better. Scores of the methods marked with * are taken from (Gupta et al. 2017).

Model	Beam size	BLEU	METEOR
Seq2Seq	1	29.9	24.7
VAE-SVG*	1	39.2	29.2
VAE-SVG-eq*	1	37.3	28.5
Our method	1	**40.3**	**30.1**
Seq2Seq*	10	33.4	25.2
Residual LSTM*	10	37.0	27.0
VAE-SVG*	10	41.3	30.9
VAE-SVG-eq*	10	39.6	30.2
Our method	10	**42.6**	**31.3**

Table 2: Results on Quora dataset. Higher BLEU and METEOR score is better. Scores of the methods marked with * are taken from (Gupta et al. 2017).

Model	Beam size	BLEU	METEOR
Seq2Seq	1	25.9	25.8
Residual LSTM	1	26.3	26.2
VAE-SVG*	1	25.0	25.1
VAE-SVG-eq*	1	26.2	25.7
Our method	1	**27.6**	**29.9**
Seq2Seq	10	27.9	29.3
Residual LSTM	10	27.4	28.9
VAE-SVG-eq*	10	37.1	32.0
Our method	10	28.4	30.6

Figure 3: Visualization of dictionary-guided attention in the decoder. Each column in the diagram corresponds to the weights of the decoder and items in the paraphrased dictionary.

VAE-SVG is the SoTA of Quora, while VAE-SVG-sq is the SoTA of MSCOCO

Case Study

- red denotes paraphrased dictionary pairs
- blue denotes phrases which are found in the paraphrased dictionary

Source	these two cats are playing in a room that has a large tv and a laptop computer .
Reference	a cat being lazy and a cat being nozy in a living room with tv and a laptop displaying the same things .
Generated	two cats are playing in a living room with a television .
Dictionary	(a tv, a television) (a large and, a great and) (playing a, to play a)

Source	a large passenger airplane flying through the air .
Reference	an airplane that is , either , landing or just taking off .
Generated	a large jetliner flying through a blue sky .
Dictionary	(the airplane, the aeroplane) (airplane, jetliner) (a large, a great)

Source	a large passenger airplane flying through the air .
Reference	an airplane that is , either , landing or just taking off .
Generated	a large jetliner flying through a blue sky .
Dictionary	(the airplane, the aeroplane) (airplane, jetliner) (a large, a great)

Source	what are ways i can make money online ?
Reference	can i earn money online ?
Generated	how can i earn money online easily ?
Dictionary	(make money, earn money) (indicate in what ways, indicate how) (making money, earn money)

Source	can you offer me any advice on how to lose weight ?
Reference	how can i efficiently lose weight ?
Generated	can you give me some advice on losing weight ?
Dictionary	(offer advice, provide advice) (offer advice, give advice) (you lost weight, you 've lost weight)

Conclusion

- The system **jointly** learns the selection of the appropriate word level and phrase level paraphrase pairs in the context of the original sentence from the Paraphrase Database (PPDB)

- The proposed model is able to replace some words or phrases in the original sentence based on the dictionary and **makes necessary changes** to ensure the new sentence is grammatically correct and fluent

END & Thanks