Story-level Text Style Transfer: A Proposal

Yusu Qian
Computer Science Department
Courant Institute of Mathematical Sciences
New York University
251 Mercer St 801, New York, NY 10012
`yq729@nyu.edu`

Abstract

Text style transfer aims to change the style of the input text to the target style while preserving the content to some extent. Previous works on this task are on the sentence level. We aim to work on story-level text style transfer to generate stories that preserve the plot of the input story while exhibiting a strong target style. The challenge in this task compared to previous work is that the structure of the input story, consisting of leading roles and their relations with each other, needs to be preserved, and that the generated story needs to be consistent after adding flavors. We plan to explore three methods including the BERT-based method, the Story Realization method, and the Graph-based method.

1 Introduction

Text style transfer has been extensively explored by the NLP community on the sentence level. In previous work, researchers defined style of a sentence as one or some of its attributes, including but not limited to sentiment (Xu et al., 2018; John et al., 2019; Liao et al., 2018), formality (Luo et al., 2019; Jain et al., 2018; Rao and Tetreault, 2018), factuality (Zhang et al., 2018), etc. The goal is to change the specified attribute or attributes in the input sentence to the target attribute or attributes. For example, changing a positive sentence to a negative sentence while keeping its key information. There are also works on transferring Shakespearean English to modern English and backward (Xu et al., 2012; Jhamtani et al., 2017).

In this paper, we propose methods to transfer text style on the story level. The task takes a story as input, and generates a story in the target style with the main plot of the input story preserved. In our work, we define *style* as the setting of the story which reveals time background and geographical information. For example, if a story starts with a boy receiving a package containing *parchments* and a *robe* delivered by an *owl*, a good guess is that this is a magic story most likely taken from or inspired by *Harry Potter*. If we want to change the above mentioned story into the *Alice in Wonderland* style, an ideal output maybe a story about a girl receiving a package containing an invitation to a tea party from a rabbit.

Compared with sentence-level text style transfer, our proposed work faces more challenges. It is impractical to collect parallel stories that have the same plot or structure but differ in settings. To deal with this, we break down the task into two parts. First, we explore methods to build a structural representation of the original story to preserve the main plot, including leading roles and their connections. Second, we generate a story given the retrieved information or graph and the target style.

2 Related Work

The work we propose is closely related to previous work on event extraction from text so that we have a structure representation of the input story, and text generation from events to produce the story in the target style.

Graph Extraction from Text Generating text on the story level from events ideally requires the events to be organized as a structural representation, otherwise the plot will not be consistent. While manually constructing graphs is expensive, there are multiple approaches to automatically construct graphs based on stories, including Named Entity Recognition (NER), Knowledge Graph, and other text graph generation methods. While NER has been studied for a while, the task of extracting named entities with semantic relations between nodes labelled has much room left to be explored. Most previous work on extracting entities together with relations either extract them separately and

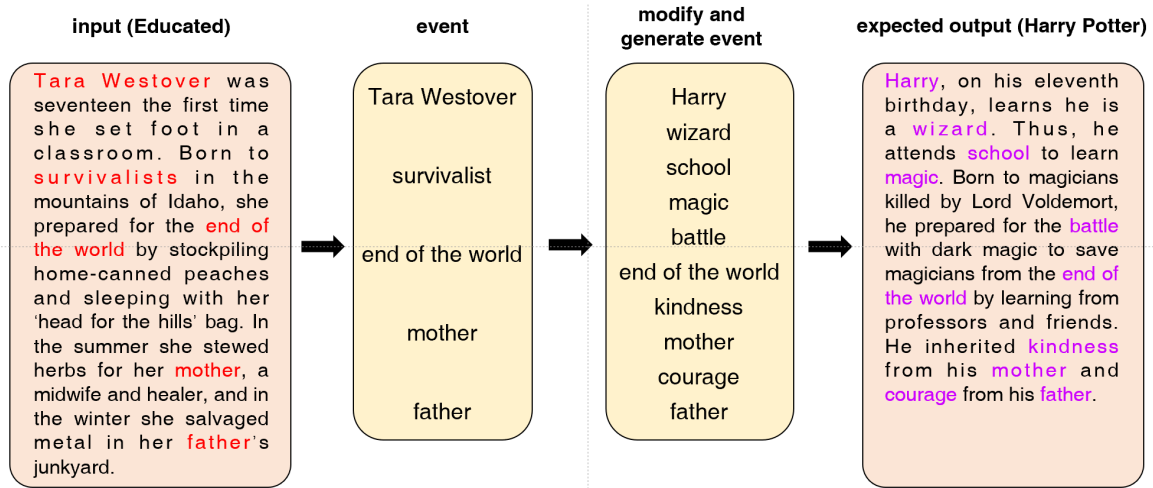| input (Educated)	event	modify and generate event	expected output (Harry Potter)
Tara Westover was seventeen the first time she set foot in a classroom. Born to survivalists in the mountains of Idaho, she prepared for the end of the world by stockpiling home-canned peaches and sleeping with her 'head for the hills' bag. In the summer she stewed herbs for her mother, a midwife and healer, and in the winter she salvaged metal in her father's junkyard.	Tara Westover survivalist end of the world mother father	Harry wizard school magic battle end of the world kindness mother courage father	Harry, on his eleventh birthday, learns he is a wizard. Thus, he attends school to learn magic. Born to magicians killed by Lord Voldemort, he prepared for the battle with dark magic to save magicians from the end of the world by learning from professors and friends. He inherited kindness from his mother and courage from his father.

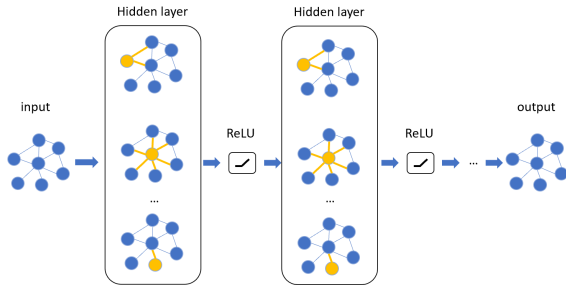Figure 1: Illustration of how Story-realization Method works

Figure 2: Structure of Graph Convolutional Network

predict the latter given the former (Chan and Roth, 2011; Zhou et al., 2005), or rely on feature engineering to extract them jointly (Ren et al., 2017).

Recently, Fu et al. (2019) employed relation-weighted graph convolutional networks (GCNs) (Kipf and Welling, 2016) to build an end-to-end relation extraction model, GraphRel, and reported SOTA results. Figure 2 illustrates how GCN works. GCN is a variant of convolutional neural networks (CNNs) that works on graphs. The representation of each node is updated based on its adjacent nodes.

Text Generation from Graph Due to the variety of graphs and information loss of long-distance dependencies, it is hard to generate coherent stories that span across multiple sentences from a graph. Koncel-Kedziorski et al. (2019) proposed a novel graph transformer to alleviate this problem by leveraging the relational structure of graphs without setting linearization or hierarchical constraints.

The usage of GCN for text generation from graphs is enjoying growing popularity among researchers. Marcheggiani and Perez-Beltrachini (2018) used GCNs to build an encoder which calculates the node representation of each node in a directed graph. After adding residual connections and dense connections between the GCN layers, they used an LSTM decoder. Guo et al. (2019) built Densely Connected Graph Convolutional Networks to address the issue of learning deeper GCNs, and achieved better results on graph-to-sequence learning and AMR-to-text generation than previous methods.

3 Proposed Methodology

Our goal is to adapt the original story to the target setting. A well-known example of such kind of adaptation is New York theatre production *Sleep No More*, which adapts the story of Macbeth deprived of its original time setting, and sets in a 1930s hotel called the McKittrick.

3.1 Data Set

The data sets ideal for our proposed work need to satisfy the following requirements. First, each corpus needs to have an abundant amount of text in the same *style*. Second, the *style* of each corpus should differ from each other significantly, to the extent that a snippet from a certain corpus tells enough for people to tell which corpus it is from.

We select paragraphs between 100 and 200 words from each corpus and use GraphRel to automatically build graphs from the text.

For each method described in the next section, we use different training data. For the BERT-based method, we use the story corpora as training data. For the Story-realization method, we use the selected paragraphs and corresponding extracted
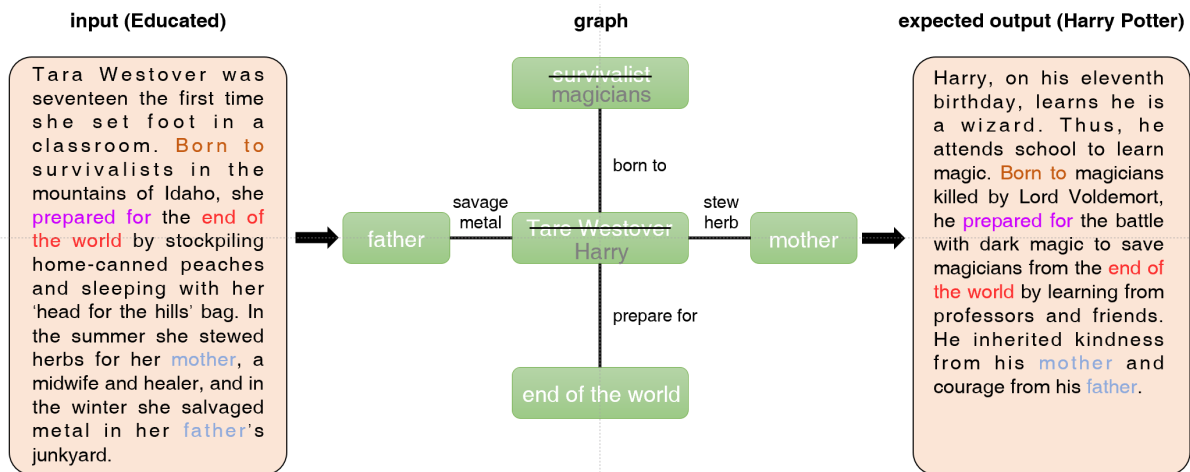
Figure 3: Illustration of how Graph-based Method works

named entities as training data. For the Graph-based method, we use the selected paragraphs and graphs built by GraphRel as training data.

To satisfy these requirements, we choose the *Harry Potter Series* and the *Game of the Throne Series* as our corpora. The former consists of 1,084,170 words and the latter consists of 1,736,054 words.

3.2 Models

We plan to experiment with the following three methods. The first two methods serve as baselines.

BERT-based Method This method will be based on Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018). In this method, first we use the corpus in the target style to fine-tune BERT. Then we build a vocabulary for the target corpus, setting the threshold of minimum occurrence to 20. We examine each word in the input story to see if they are included in the vocabulary of the target corpus. If they are not, we use the fine-tuned BERT to mask and predict these words one by one. The BERT-based method serves as our baseline model as it modifies the input story sentence by sentence instead of as a whole.

In simpler cases where we only wish to change the era of the story and do not have any other requirements, we can add append a phrase indicating the era to the original sentences. For example, when we mask *video* in the sentence 'The boy spent a whole day playing video games.', BERT (large-cased version) correctly predicts the word to be *video*. If we add the phrase 'on the first day of the 18th century', the prediction becomes *card*, which matches the time setting.

Story Realization Method Ammanabrolu et al. (2019) proposed an ensemble-based model to generate sentences given plot events. This involves two steps. First, we need to extract events from the input story. This can be done through Named Entity Recognition (in this work we will use Al-lennlp NER) and finding VerbNet (Kipper-Schuler, 2005) classes of verbs and WordNet (Miller, 1995) Synsets for nouns recognized as events. The next step is to expand these events to a story. We plan to experiment with the ensemble model by Ammanabrolu et al. (2019) which is reported to combine the strength of the retrieve-and-edit method (Hashimoto et al., 2018), the template filling method, the sequence-to-sequence methods with finite state machine decoder, Monte Carlo beam decoding, and vanilla beam-decoding respectively. This method will conduct an event-to-event generation first to include more events before generating the output story. Figure 1 illustrates how this method works.

Here we need to note that sometimes extracted entities or relations are out-of-target-corpus-vocabulary words in the target style corpus. For example, *computer* is not in the corpus of *Harry Potter*. We need to replace these words with words that have the same part of speech and closest in the word embedding trained on the target corpus. Euclidean distance is used for distance calculation.

We expect that compared with the BERT-based method, the Story Realization method will perform better in terms of creativity while not as well in terms of content preservation.

10

Graph-based Method In this method, a similar replacing scheme of out-of-target-corpus-vocabulary words as in the story realization method should be used on the input story. Then we plan to experiment with graph transformers and other graph-to-text generators trained on our data sets, compare their performance on our task, and examine the possibility to improve their performance by making modifications. Specifically, in the text-to-graph step we explore using Graph Neural Network. We plan to start with using GraphRel, the GCN-based SOTA entity and relation extraction model, to convert the input story to a graph.

Figure 3 illustrates how the Graph-based Method works. The input is an extract from the novel *Educated*. A graph containing key information is built upon the input story. Some modification is done to replace out-of-target-corpus-vocabulary words. We expect the output to preserve the structure of the input story while being creative and consistent. Towards this goal, we plan to experiment with different GCNs structures for text generation.

3.3 Evaluation

We plan to evaluate our generated stories using perplexity and human evaluation, with an emphasis on the latter considering the creative nature of this task.

The generated stories will be evaluated by linguists from these aspects: grammar and fluency; main plot preservation; strength of the target style; creativeness. Each aspect will be given a score between 1 and 5, with 1 representing total failure, 2 representing barely acceptable, 3 representing acceptable, 4 representing good, and 5 representing the most satisfying performance.

4 Summary

We propose to explore text style transfer on the story level. The challenge remains in preserving the main plot and generating consistent and meaningful text in the target style. We plan to focus mostly on studying the possible application of GCN in this task. We will perform extensive experiments and report results in future work.

5 Acknowledgment

References

Prithviraj Ammanabrolu, Ethan Tien, Wesley Cheung, Zhaochen Luo, William L. White Ma, Lara J. Martin, and Mark O. Riedl. 2019. Story realization: Expanding plot events into sentences. *ArXiv*, abs/1909.03480.

Yee Seng Chan and Dan Roth. 2011. Exploiting syntactico-semantic structures for relation extraction. In *The 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference, 19-24 June, 2011, Portland, Oregon, USA*, pages 551–560. The Association for Computer Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. BERT: pre-training of deep bidirectional transformers for language understanding. *CoRR*, abs/1810.04805.

Tsu-Jui Fu, Peng-Hsuan Li, and Wei-Yun Ma. 2019. GraphRel: Modeling text as relational graphs for joint entity and relation extraction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1409–1418, Florence, Italy. Association for Computational Linguistics.

Zhijiang Guo, Yan Zhang, Zhiyang Teng, and Wei Lu. 2019. Densely connected graph convolutional networks for graph-to-sequence learning. *Transactions of the Association for Computational Linguistics*, 7:297–312.

Tatsunori B. Hashimoto, Kelvin Guu, Yonatan Oren, and Percy Liang. 2018. A retrieve-and-edit framework for predicting structured outputs. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS'18, page 10073–10083, Red Hook, NY, USA. Curran Associates Inc.

Parag Jain, Abhijit Mishra, Amar Prakash Azad, and Karthik Sankaranarayanan. 2018. Unsupervised controllable text formalization. *CoRR*, abs/1809.04556.

Harsh Jhamtani, Varun Gangal, Eduard Hovy, and Eric Nyberg. 2017. Shakespearizing modern language using copy-enriched sequence-to-sequence models. *EMNLP 2017*, 6:10.

Vineet John, Lili Mou, Hareesh Bahuleyan, and Olga Vechtomova. 2019. Disentangled representation learning for non-parallel text style transfer. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 424–434, Florence, Italy. Association for Computational Linguistics.

Thomas Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *ArXiv*, abs/1609.02907.

Karin Kipper-Schuler. 2005. *VerbNet: a broad-coverage, comprehensive verb lexicon*. Ph.D. thesis, Computer and Information Science Department, Universiy of Pennsylvania, Philadelphia, PA.

Rik Koncel-Kedziorski, Dhanush Bekal, Yi Luan, Mirella Lapata, and Hannaneh Hajishirzi. 2019. Text Generation from Knowledge Graphs with Graph Transformers. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 2284–2293, Minneapolis, Minnesota. Association for Computational Linguistics.

Yi Liao, Lidong Bing, Piji Li, Shuming Shi, Wai Lam, and Tong Zhang. 2018. QuaSE: Sequence editing under quantifiable guidance. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3855–3864, Brussels, Belgium. Association for Computational Linguistics.

Fuli Luo, Peng Li, Jie Zhou, Pengcheng Yang, Baobao Chang, Xu Sun, and Zhifang Sui. 2019. A dual reinforcement learning framework for unsupervised text style transfer. In *IJCAI*.

Diego Marcheggiani and Laura Perez-Beltrachini. 2018. Deep graph convolutional encoders for structured data to text generation. In *Proceedings of the 11th International Conference on Natural Language Generation*, pages 1–9, Tilburg University, The Netherlands. Association for Computational Linguistics.

George A. Miller. 1995. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41.

Sudha Rao and Joel R. Tetreault. 2018. Dear sir or madam, may i introduce the gyafc dataset: Corpus, benchmarks and metrics for formality style transfer. In *NAACL-HLT*.

Xiang Ren, Zeqiu Wu, Wenqi He, Meng Qu, Clare R. Voss, Heng Ji, Tarek F. Abdelzaher, and Jiawei Han. 2017. Cotype. *Proceedings of the 26th International Conference on World Wide Web - WWW '17*.

Jingjing Xu, Xu Sun, Qi Zeng, Xiaodong Zhang, Xuancheng Ren, Houfeng Wang, and Wenjie Li. 2018. Unpaired sentiment-to-sentiment translation: A cycled reinforcement learning approach. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 979–988, Melbourne, Australia. Association for Computational Linguistics.

Wei Xu, Alan Ritter, Bill Dolan, Ralph Grishman, and Colin Cherry. 2012. Paraphrasing for style. In *Proceedings of COLING 2012*, pages 2899–2914, Mumbai, India. The COLING 2012 Organizing Committee.

Zhirui Zhang, Shuo Ren, Shujie Liu, Jianyong Wang, Peng Chen, Mu Li, Ming Zhou, and Enhong Chen. 2018. Style transfer as unsupervised machine translation.

GuoDong Zhou, Jian Su, Jie Zhang, and Min Zhang. 2005. Exploring various knowledge in relation extraction. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*, pages 427–434, Ann Arbor, Michigan. Association for Computational Linguistics.