

# Synthesizing Description of Knowledge Graphs through Text Generation

Aakrit Singhal\*

\*Podar International School, Mumbai - 400054, India

DOI: 10.29322/IJSRP.12.01.2022.p12103  
<http://dx.doi.org/10.29322/IJSRP.12.01.2022.p12103>

**Abstract-** Text style transfer aims to change the style of the input text in graphs to the target style while preserving the content to some extent. In previous work, researchers defined style of a sentence as one or some of its attributes, including but not limited to sentiment, formality, factuality, etc. Hence, the goal has been 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. However, due to the lack of parallel data and evaluation metric, there has been a slow growth in this field compared to others. To explore this, the paper includes several current experiments and pipelines that have been modified for the purpose. Datasets such as the Yelp Reviews have been used, along with other sentiment analysis dataset, and modified setup and code to get original results and get an understanding on the existing solutions.

**Index Terms-** deep learning, natural language processing, knowledge graphs, text generation, style transfer

## I. INTRODUCTION

Text generation has been a computationally complex task as it involved generating texts with complex semantical understandings of content and context of language. A recent paper [1] proposes to solve the problem of generating coherent multi-sentence texts from the output of knowledge graphs, an information extraction system. As we know, language generation consists of 2 broad components – realization and planning. Realization is generating text regardless of its meaning, relation to other texts, or overall semantics. However, planning is the process of ensuring long-term dependencies between entities in the text are resolved and are related to each semantically. The Graph Writer generates an abstract from words in the title and the constructed Knowledge graph. This paper has proposed a new graph transforming encoder which improves the relation structure of knowledge graphs by applying sequence transformers to graphs structured inputs, without imposing linearization constraints.

Similar to text generation, text-style generation is changing the style of text in a knowledge graph whilst preserving its contents. Recent papers propose [2][5][6] to improve content preservation during style transfer to minimize the content changing after text style transfer. This is done by using an Attentional auto encoder and a binary style classifier to generate sentences within the target style. Then it uses the distance between POS information of the input and output sentence. Most of the content of sentences is

captured by nouns, hence the model forces the decoder to output sentences with similar nouns (this is different for other models, where some might use verbs as better information).

A POS Tagger is used to extract nouns from both input and generated sentences, while a content conditional language model to enforce the output sentence to contain desired nouns, as well as make the output sentence more fluent.

## II. RELATED WORKS

### A. Text Generation

For Text Generation, Koncel-Kedziorski et al. [1] created the AGENDA dataset, with 40k paper titles and abstracts from semantic scholar corpus taken from proceedings of 12 AI conferences. The dataset was fed to the model after the graph creation and pre-processing.

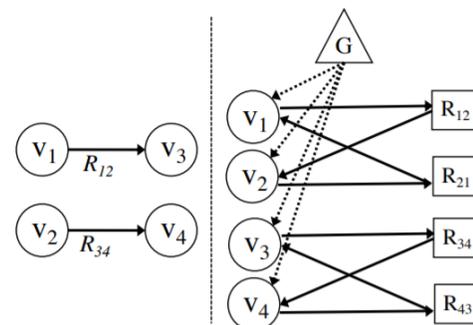


Figure 1 [1]: Converting disconnected labelled graph to connected unlabeled graph for use in attention-based encoder.

As shown in the above picture, after the SciIE system of Luan et al. [3], a state-of-the-art science main information extraction system is applied, it extracts entities and labels them, and also creates coreference annotations, which are then converted to single labelled edges. Each graph is then converted to a connected graph because of an added Global Node that all other nodes are connected to represented by  $G$  in Figure 1.

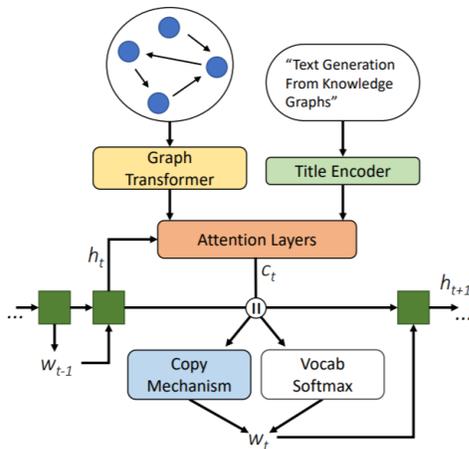


Figure 2 [1]: GraphWriter Model Overview

In Figure 2, the title is encoded with a BiLSTM [8], using 512-dimensional embeddings and no pre-trained word embeddings were used for this.

For the graph encoder, each vertex representation  $v_i$  is contextualized by attending over the other vertices to which  $v_i$  is connected in  $G$ . The graph encoder then concatenates the products of attention weights resulting from  $N$  attention heads to create a set of vertex embeddings. These embeddings are then augmented with block networks consisting multi-headed attention layers, with block of two layers each. Multiple blocks help information to propagate through the graph. The resulting vertex encodings represent entities, relations, and the global node contextualized by their relationships in the graph structure.

The decoder was attention based, with a copy mechanism for copying input from the knowledge graph and title. Moreover, a hidden state  $h_t$  is used at each time step  $t$ . From the encodings created with the encoder, context vectors  $C_g$  and  $C_s$  are computed using multi-headed attention. These context vectors are then concatenated together to make  $C_t$  the total context vector. From this, the probability of copying a token from a word is calculated, and then the final next token probability distribution.

### A. Text Style Transfer

For Text Style transfer, Tian et al. [2] experimented on sentiment manipulation that is investigating if the model can maintain the sentiment-independent content and prevent it from changing. A filtered form of the Yelp reviews dataset is used and tokens are added for certain situations to sentences. Accuracy and BLEU score is then calculated and also the POS information to test model's content preserving ability. A human annotated sentence dataset is also used that includes human generated sentences with desired style and preserved content. Hence, BLEU scores against both original, and human sentences are used to evaluate the models.

Fu et al. [4] also proposes style transfer through non-parallel data due to the lack of availability of parallel data when attempting to transfer styles. The paper aims to two evaluation metrics to measure to style transfer aspects: transfer strength and content

preservation. For this, two transfer tasks are used: Paper-news title transfer and positive-negative review transfer. The paper proves that the evaluation metrics proposed highly correlate with the human judgement.

Here, the first task is the transferring of source style to target style, where transfer strength evaluates whether the style is transferred. There were more than 100,000 training data for this task, and an LSTM-sigmoid classifier was used which performs well in big data. Here transfer strength was defined as  $\frac{N_{right}}{N_{total}}$  wherein  $N$  is the number of test data right, upon the total test data. Here, the target style was positive when the output  $\leq 0.5$ , and negative when output  $> 0.5$ .

The second task, content preservation, evaluated the similarity between source text and target text. Content preservation rate is defined as cosine distance between source sentence embedding  $v_s$  and target sentence embedding  $v_t$ . For this, pretrained 100-dimensional word embeddings were used to compute sentence similarities. For the positive and negative review transfer task, sentiment words were filtered out the to make sure the content preservation metric indeed measured the content similarity. A positive and negative word dictionary is used to conduct the filtering.

## III. RESULTS

The code implementation used an imported modified dataset with reduced size. The required libraries were installed according to stated dependencies. Then the train.py file ran, which generate a log file of the trained weights after 20 epochs with learning rate of 0.1. This was the last epoch trained:

```

14148 of like 40k -- current avg loss 2.8695265429822747
15748 of like 40k -- current avg loss 2.8878465237882397
17348 of like 40k -- current avg loss 2.900720888715695
18948 of like 40k -- current avg loss 2.911327529109579
20548 of like 40k -- current avg loss 2.9211453284552666
22148 of like 40k -- current avg loss 2.9307954230419626
23748 of like 40k -- current avg loss 2.9387807161373907
25348 of like 40k -- current avg loss 2.9454583353813426
26948 of like 40k -- current avg loss 2.9517737993133055
28548 of like 40k -- current avg loss 2.9579111340501254
30148 of like 40k -- current avg loss 2.962266869960642
31748 of like 40k -- current avg loss 2.9664896243432914
33348 of like 40k -- current avg loss 2.970494030876613
34948 of like 40k -- current avg loss 2.9740522094458917
36548 of like 40k -- current avg loss 2.977762474538305
38148 of like 40k -- current avg loss 2.9811194496508775
AVG TRAIN LOSS: 2.9819427898846382 PPL: 19.72610311503156
Evaluating in present a 0 for on 0 to with I the the can wit
VAL LOSS: 3.5793088054656983 PPL: 35.84875381901921
Saving model
    
```

Figure 3: Last Epoch training details

Then used the file of the last loss (vloss19) to run generator.py to generate 1000 sentences, with the corrected sentence structure. Below are the first 4 outputs I was able to generate:

```

0 1000
/content/drive/My Drive/GraphWriter/models/newmodel.py:152: UserWarning: To copy construct from a tensor, it is recommended to use sourceTensor.clone_
  ex = torch.tensor(hv)
we present a LEARNING ARCHITECTURE for LEXICAL SEMANTIC CLASSIFICATION PROBLEMS that <unk> TASK-SPECIFIC TRAINING DATA with BACKGROUND DATA encoding
in this paper , we present a novel LEARNING ARCHITECTURE for LEXICAL SEMANTIC CLASSIFICATION PROBLEMS in LEXICAL SEMANTIC CLASSIFICATION PROBLEMS .

1 1000
we consider the STOCHASTIC MULTI-ARMED BANDIT PROBLEM with a PRIOR DISTRIBUTION on the REWARD DISTRIBUTIONS . we are interested in studying PRIOR-FR
in this paper , we study the STOCHASTIC MULTI-ARMED BANDIT PROBLEM in the STOCHASTIC MULTI-ARMED BANDIT PROBLEM with REWARD DISTRIBUTIONS . the STOCH

2 1000
the paper presents a new application of AUTOMATIC SPEECH PROCESSING in the AMBIENT ASSISTED LIVING AREA , developed in the course of a three year re
in this paper , we describe the development of AUTOMATIC SPEECH PROCESSING and SPEAKER TRACKING in the context of ELDERLY CONVERSATIONAL SPEECH . the

3 1000
we address the problem of estimating the WORD ERROR RATE of an AUTOMATIC SPEECH RECOGNITION SYSTEM without using ACOUSTIC TEST DATA . this is an imp
in this paper , we propose an AUTOMATIC SPEECH RECOGNITION SYSTEM based on LEAVE-ONE-OUT DECODING and GRAMMAR DESIGN . the AUTOMATIC SPEECH RECOGNIT

4 1000
mel-frequency cepstral coefficients -lrb- MEL-FREQUENCY CEPSTRAL COEFFICIENTS -rrb- are the most widely used FEATURES for SPEECH RECOGNITION . these
in this paper , we propose a novel method for SPEECH RECOGNITION based on MEL-FREQUENCY CEPSTRAL COEFFICIENTS . the proposed method is based on the
    
```

Figure 4: First four generator.py outputs of abstracts

Below is an example from the output:

Original Abstract:

this paper proposes an approach for AUTOMATIC ROAD EXTRACTION in AERIAL IMAGERY which exploits the SCALE-SPACE BEHAVIOR OF ROADS in combination with GEOMETRIC CONSTRAINED SNAKE-BASED EDGE EXTRACTION . the approach not only has few parameters to be adjusted , but for the first time allows for a BRIDGING OF SHADOWS and partially occluded areas using the heavily disturbed evidence in the image . the road network is constructed after extracting <unk> of various shape and topology . reasonable results are obtained which are evaluated based on ground truth.

The Graph Writer Generated Abstract:

in this paper , we present a method for AUTOMATIC ROAD EXTRACTION in AERIAL IMAGERY . the approach is based on GEOMETRIC CONSTRAINED SNAKE-BASED EDGE EXTRACTION from AERIAL IMAGERY . the algorithm is based on GEOMETRIC CONSTRAINED SNAKE-BASED EDGE EXTRACTION in AERIAL IMAGERY . the algorithm is based on GEOMETRIC CONSTRAINED SNAKE-BASED EDGE EXTRACTION and is shown to be robust against AERIAL IMAGERY . the algorithm has been tested on a variety of AERIAL IMAGERY , and it is shown that the proposed method has better performance than existing methods .

When first trained on all columns of the test.tsv dataset and ran the eval.py file, the following BLUE scores were received:

Table 1: Scores for running entire script

<b>Bleu_1</b>	21.056712675994802
<b>Bleu_2</b>	12.158468404347657
<b>Bleu_3</b>	7.304507502503695
<b>Bleu_4</b>	4.484571389432444
<b>ROUGE_L</b>	16.145650769074155

However, these results weren't matching with the expected results, and hence only the Golden Targets in test.tsv were used. Hence, when trained on that, results has a higher accuracy to the original.

Table 2: Scores for running selected column

<b>Bleu_1</b>	31.306484663928753
<b>Bleu_2</b>	13.864942960995666
<b>Bleu_3</b>	6.704610880358337
<b>Bleu_4</b>	3.7343222819736184
<b>ROUGE_L</b>	20.172933649668305

Blue\_2 matched was the closest to the result on the paper, which was  $14.3 \pm 1.01$ .

#### IV. CONCLUSION

The paper proposes to generate texts with complex accurate semantic fields accurately. It also expands on changing style of the text while preserving its contents. Possible extensions could be to use different languages for text generation, and perform style transfer without parallel data. This can also be helpful in generating large quantities of text that are randomized and semantically not accurate in nature.

#### REFERENCES

- [1] R. Koncel-Kedziorski, D. Bekal, Y. Luan, M. Lapata, and H. Hajishirzi, "Text Generation from Knowledge Graphs with Graph Transformers," *Proceedings of the 2019 Conference of the North*, 2019.
- [2] Y. Tian, Z. Hu, and Z. Yu, "Structured Content Preservation for Unsupervised Text Style Transfer," *arXiv*, no. 1810.06526, 2018.
- [3] Y. Luan, L. He, M. Ostendorf, and H. Hajishirzi, "Multi-task identification of entities, relations, and coreference for scientific knowledge graph construction," *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018.
- [4] Z. Fu, X. Tan, N. Peng, D. Zhao, and R. Yan, "Style Transfer in Text: Exploration and Evaluation", *AAAI*, vol. 32, no. 1, Apr. 2018.
- [5] Y. Qian, "Story-level text style transfer: A proposal," *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop*, 2020.
- [6] Y. Shi, S. Zhang, C. Zhou, X. Liang, X. Yang, and L. Lin, "GTAE: Graph Transformer-based auto-encoders for linguistic-constrained text style transfer," *ACM Transactions on Intelligent Systems and Technology*, vol. 12, no. 3, pp. 1–16, 2021.
- [7] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [8] M. Schmitt, S. Sharifzadeh, V. Tresp, and H. Schütze, "An unsupervised joint system for text generation from knowledge graphs and semantic parsing," *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020.
- [9] J. Li, T. Tang, W. X. Zhao, Z. Wei, N. J. Yuan, and J.-R. Wen, "Few-shot knowledge graph-to-text generation with pretrained language models," *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 2021.

#### AUTHORS

**First Author** – Aakrit Singhal, Student, Podar International School, [aakritsinghal44@gmail.com](mailto:aakritsinghal44@gmail.com)