Keyword-Based Control in Story Generation

AKSHAYA J PATIL, Department of CSE, Penn State University amp8136@psu.edu

Additional Key Words and Phrases: natural language processing, text generation
Akshaya J Patil. . Keyword-Based Control in Story Generation.

1 INTRODUCTION

Story generation is an application of natural language processing (NLP) that aims to generate coherent and engaging stories automatically. Generating text that accurately reflects a user's intended meaning is a significant challenge in text generation. Several approaches that enable users to manually specify the keywords or phrases that should be included in the generated text have been developed to address this issue.

While providing keywords or phrases can be a useful way to guide the generation of text, it may not always be sufficient to ensure that the generated text reflects the user's intent in an accurate way. To ensure that the keywords are given the right amount of significance, it may occasionally be required to reposition the keywords inside the created text. In this project using the technique of adjusting the position of keywords within a generated text is applied to ensure the generated text accurately reflects the user's intent, particularly in cases where the intended meaning of the keywords is complex.

2 RELATED WORK

2.1 Controllable Abstractive Summarization [1]

The goal of the paper is to create a brief overview that encapsulates the essential details of a lengthy work. The authors suggest a new architecture that gives users flexibility over the summary's level of abstraction and scope. This is accomplished by including two new mechanisms: a copy mechanism that allows the model to only copy specific pieces of data from the input document, and a conditional variational autoencoder that creates a summary based on the user's input.

2.2 Plan-and-Write: Towards Better Automatic Storytelling [2]

Planning and generation are the two stages of the independent storytelling method used in this paper. An outline of the main characters, events, and relationships in the novel is created during the planning stage. The plot is written based on the plan during the generation stage. It produces plans and tales jointly using a neural network model. A seq2seq model with bidirectional gated recurrent units (BiGRUs) that first predicts the plot from the title (as indicated by a series of keywords). The story is then generated based on both the title and the plot.

3 DATA

ROC Stories[3] is a publicly available dataset of fictional narratives aimed at advancing research in natural language processing (NLP), particularly in the area of story generation. The dataset consists of 98,161 short stories written by authors on the internet-based writing platform Reddit. Each story in the dataset comprises five sentences. These stories are typically several paragraphs long and are

Author's address: Akshaya J Patil, Department of CSE, Penn State University amp8136@psu.edu.

Dataset size	Total Data	Training Data	Validation Data	Testing Data
ROC_stories	98161	78,528	9,816	9,816

Table 1. Dataset statistics

written in a variety of genres, including science fiction, romance, horror, and fantasy. The data is split into 8:1:1 for training, validation, and testing datasets.

4 IMPLEMENTATION AND REASONING

For the task of story generation, we use the GPT model [4]. GPT-2 is a transformer-based language model that has been pre-trained on a massive corpus of text and has shown remarkable performance on various natural language processing tasks. It has a large number of parameters, which enables it to capture complex patterns and relationships in the input data, making it well-suited for text generation tasks such as story generation. Additionally, GPT-2 has the ability to generate coherent and diverse text, which is essential for story generation where a high degree of creativity and variation is required.

In the case of the GPT model, the control tokens are fed to the decoder. The model is then trained to maximize the conditional probabilities $p(yi|y<i, x)$ by using cross-entropy loss. Here, y refers to the target text and x refers to the input to the model, including the control tokens and the source document in the case of generation.[5]

In order to reflect more particular user intentions and produce writings that pique readers' interest, it can be difficult to control the precise placement of keywords in the created content. The idea of controlling text attributes by using special tokens has been previously explored in works such as [6]. In this paper, we also use special tokens to provide the model with information on the position of each keyword and the desired length of the generated text.

Since the GPT model decoder is responsible for generating the output sequence given the input sequence. By providing the keywords to the decoder, the model can ensure that the generated text contains the desired information related to the keywords. This allows for greater control over the output text and can help ensure that the generated text is more relevant and useful to the end user. Additionally, providing the keywords to the decoder allows the model to incorporate them into the generation process, taking into account their position and importance in the final text. **Model Design**

5 EXPERIMENTS AND RESULTS

For a story generation task, we used the ROCStories dataset and the GPT2 model. Control tokens (keywords, each keyword position, and the text length) were extracted from the target text and given to the model for training. Word tokenization of the text was done by NLTK library [7] to obtain keywords and text length.

By giving the model a particular token that reflects the desired relative location of the keyword (for example, 0–10%) and desired text length, the experiment is performed. The task of creating stories The learning rate is set to $2x10^{-5}$, while the word embedding weights' learning rate was set to $1x10^{-3}$. There are five epochs. Texts were produced using top-p sampling, where p = 0.95. The thermostat was set to 0.1 degrees. The target text's maximum token count was set to 128, and any

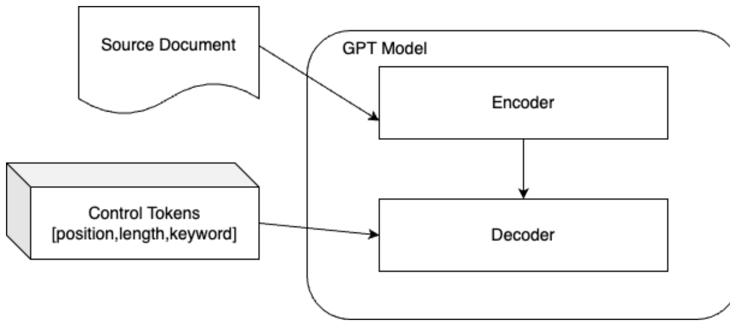

Fig. 1. The model architecture overview

text with more tokens was terminated at the end.

Keyword-Based Control Approach By providing the model with all control tokens during training and generating texts under 2 keywords consideration (Keyword, +Len+Pos), we can use the same model for multiple experiments without the need to train a separate model for each setting. This reduces the overall cost and time required for experimentation. The reason for choosing this approach is to enable a single model to perform inference experiments in multiple settings with lower costs.[5] We evaluate the accuracy of generating text including all target keywords and the accuracy of generating text in which all target keywords are placed in each target position. Table 2 displays the accuracy of creating text that contains every target keyword as well as the accuracy of generating text that uses every target keyword in every target position.

Control	Including	POS
No keyword	0.5	0.1
Keyword	19.7	1.2
+POS+Len	23	6.3

Table 2. Evaluation of the accuracy of with and without control of keywords size-2 and their positions

6 ANALYZE

The accuracy of the keyword inclusion and the keyword position control is low, especially in story generation. The reason for maybe that the model does not generate the appropriate context for the inclusion of keywords because the source document is not given.

REFERENCES

[1] Angela Fan, David Grangier, and Michael Auli. 2018. Controllable Abstractive Summarization. In Proceedings of the 2nd Workshop on Neural Machine Translation and Generation, pages 45–54, Melbourne, Australia. Association for Computational Linguistics.

[2] Yao, L., Peng, N., Weischedel, R., Knight, K., Zhao, D., Yan, R. (2019). Plan-and-Write: Towards Better Automatic Storytelling. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 7378-7385.

[3] Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A corpus and cloze evaluation for deeper understanding of commonsense stories. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 839–849, San Diego, California. Association for Computational Linguistics.

[4] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2018. Language models are unsupervised multitask learners.

[5] Sasazawa, Y., Morishita, T., Ozaki, H., Imaichi, O., Sogawa, Y. (2023). Controlling keywords and their positions in text generation. arXiv preprint arXiv:2304.09516.
[6] Kango Iwama and Yoshinobu Kano. 2019. Multiple news headlines generation using page metadata. In Proceedings of the 12th International Conference on Natural Language Generation, pages 101–105, Tokyo, Japan. Association for Computational Linguistics.
[7] https://github.com/nltk/nltk/blob/develop/nltk/tokenize