# Natural sentence generation from the knowledge graph using the Transformer neural network machine learning approach

Jiachen Li
1068299

## Introduction

Natural language generation (NLG) was a sub-problem of the natural language processing (NLP) problem. NLG aimed to generate human-readable natural language from knowledge bases. Knowledge bases were the technologies used to store abundant information on the computer in a structured way designed to be easily accessed by computer systems. However, knowledge bases have been challenging for humans to consume and utilise before converting to natural sentences. The NLG field of research has existed for a long time because it plays an essential role in areas such as Artificial Intelligence or question-answering systems. But it has gained significantly more attention recently due to the popularity of Metaverse[1].

The Metaverse has been attracting an excessive amount of attention in recent years. Its popularity was due to the potential market value, which is expected to have a compound annual growth rate of 43.20% between 2022 to 2030, and eventually reach a spectacular 39.25 billion USD market value in 2030 [2]. The NLG technology was acting as the foundation of the next-generation human-computer interactive question-answering system to provide sufficient immersion for the Metaverse.

In the early era of the studies, most proposed solutions were domain-specific rule-based systems; for example, one model was for the cooking field only [3]. However, this approach was suboptimal as it cannot generalise and require extensive human work. With that in mind, the state-of-the-art solution took advantage of the neural network encoder-decoder-based approach named the GTR-LSTM model by Trisedya, Qi, and Zhang outperformed all previous models [4]. Even though the neural approach was a giant leap from previous models, there was still room for improvement.

After initiating a new research field with GTR-LSTM, Trisedya and his team continuously improved its performance. Firstly, they flipped the problem and extracted a knowledge base from the natural language [5] to extend the use case and expand the training pool. Secondly, they studied entity alignment to match the same real-world entity from different knowledge bases [6] to merge and expand the network of knowledge bases. Moreover, Zhang has proposed a novel benchmark framework to evaluate various entity alignment strategies [7]. Lastly, they integrate content planning in the attention model [8] for more fluent natural language with correct entity order [9]–[11]. This paper is aimed to follow up on the progress of this study field and further improve the performance of deep learning-based NLG technic.

## Literature review

Before exploring proposed solutions to the research question, this paragraph will explain the problem to the general audience with limited field knowledge. The knowledge base was a class of strategies for storing information, knowledge, and truth on the computer system. In contrast, the knowledge graph was the most dominant instance of knowledge bases. It was constructed by Resource Description Framework (RDF) triple in format <subject, predicate, object>. For example, a triple would be: <Alice, residence, Melbourne>, and in the context of NLG, the generated sentence could be: “Alice lives in Melbourne”. In the real world, two commonly used knowledge graphics that will be mentioned in a later chapter of this paper is Wikidata [12] and Freebase [13].

Trisedya1 et al. proposed GTR-LSTM, a triple encoder to be used in an encoder-decoder structured neural network-based NLG solution [4] to solve the research problem.
Trisedya has shown that the neural network approach can generate more fluent and accurate sentences and generalise to all fields and requires less human effort by comparing it with classical lexicon-based NLG solution specifically for the cooking area proposed by Cimiano et al. [3]. Trisedya has also discussed other solutions besides one by Cimiano to show that the current neural-based state of art solution has been a significant step forward in solving the NLG problem and completely reshaped the field of research with new issues that have been brought up. The study aimed to better capture the relationship between subject and object within a triple (intra-triple relationships) and entities between triples within the same knowledge graph (inter-triple relationships). To better illustrate the approach, continue on

<Alice, residence, Melbourne> example mentioned before but with another triple: <Melbouren, city, Australia> within the same knowledge graph. Question: “Which country does Alice live in?” would be answered with “Alice lives in Australia.” Under the proposed approach. Note that the model found Melbourne was not the name of a country, but Australia was; also, Melbourne is a City in Australia, rather than the other way around. The RDF pre-processor module was used to achieve this feature. It mapped the entities to their type, in this example, Alice to person, Melbourne to city and Australia to Country. The model could preserve input information structure even if there were cycles or non-predefined relationships between entities in the knowledge graph. As a result, the GTR-LSTM have improved over the basic neural network encoder-decoder-based BLSTM [14] by up to 17.6%, 6.0%, and 16.4% in terms of BLEU [15], METEOR [16], and TER [17] scores.

The next step Trisedyl’s team [5] took was to utilise their knowledge and experience in developing GTR-LSTM to extract knowledge base from a natural sentence which is the opposite of what they were doing. We reviewed this paper because the natural language is more abundant than the knowledge graph, especially recently developed knowledge. By gaining the ability to extract RDF tuple from natural language, we could obtain extra labelled training sets for supervised learning. By converting information back and forth, we could better evaluate the performance and weakness of our model. The paper was structured to discuss the issues of current relation extraction technologies and propose corresponding solutions. Previous studies relied on Named Entity Disambiguation (NED) to handle the task of mapping triple to knowledge base space which was the latter half of the problem and only force on improving the first half, which is information extraction from the sentence. To address the error propagation problem in previous studies caused by NED, a novel neural-based end-to-end relation extraction model for knowledge extraction was proposed. To avoid multi-word entity names in sentences being broken apart during the extraction phase, an n-gram-based attention model was proposed. And lastly, a modified beam search and a triple classification algorithm were proposed to help generate a more precise knowledge graph by correctly merging real-world entities with different names. This could avoid the defect of traditional embedding similarity-based greedy search algorithms connecting similar but distinct entities.

As mentioned before, the knowledge graph has been the most popular form of knowledge base.  Therefore, different companies and originations have constructed their own knowledge graphics. However, different knowledge graphics could dramatically differ in how entities were represented in RDF triples. Besides, even though most information was present in most knowledge graphs, the slight extra details in some knowledge graphs could add up to a qualitative difference. To connect different knowledge graphs, entity alignment was studied. The fundamental idea was to identify and align the same real-world entity in different knowledge graphs, although they would have been represented in a different way. To help the demonstration, we defined there were two types of RDF triples that capture an entity's attribute or relation. Bring back the Alice example again; we believe triple <Alice, residence, Melbourne> was explaining Alice has an attribute that is she lives in Melbourne. On the other side, triple <Melbouren, city, Australia> has captured the inclusion relation of two geographic entities. Early studies on entity alignment were primarily based on the similarity between attributes of entities [18]. This approach was error-prone because different types of entities could have completely different attributes and it is up to the human user to make rules on which attribute is more valuable than others. The residence in the Alice example was clearly a less efficient attribute to do entity alignment because other large numbers of people could also live in Melbourne. A more recent approach was based on embedding [19]. It was done by using a transition matrix to map the embedding space of two knowledge graphs together. Unfortunately, the pre-aligned entities referred as seed alignments used to compute this matrix were not commonly available and thus required extensive human effort. Trisedya et al. [6] have proposed a new embedding model that took strength from two previous models while avoiding the drawbacks to address the above challenges. The model first generates attribute embeddings from the attribute triples and then uses them as the seed alignment to shift the entity embeddings of two knowledge graphs to the same vector space. Moreover, the transitivity rule was applied to further enrich the number of attributes of an entity to better identify the similarity between entity embeddings. Trisedya claims that the proposed model has a 50% improvement of *hits@1* over the baseline on three real-world knowledge graphs.

While researching entity alignment technics, Trisedya and his team realised the limitation in current existing benchmark frameworks for entity alignment. Also, the current research community on entity alignment was small but growing fast as NLG became more and more popular. Therefore, Zhang et al. [7] have written a paper that provided a comprehensive tutorial-type review to help researchers who are new to this research field to catch up with the development of entity alignment technics without the need to find and read an excessive number of individual papers. They also designed new benchmark datasets to solve the limitations of currently available benchmark datasets, such as bijection, lack of name variety, and small scale. The proposed benchmark dataset was named DWY-NB, where the NB stands for non-bijection. By conducting extensive experiments testing state-of-the-art entity alignment techniques using DWY-NB data, they found that models that exploit semantic information such as attribute triples tend to have significantly higher accuracy which was exactly what the model discussed in the last paragraph did.

The last technic used to improve NLG performance was content-planing. The machine learning model used for the neural NLG task was LSTM which was originally designed with natural language input in mind, so the input was consumed in a linear sequence. This was suboptimal when the input was a knowledge graph as it could not preserve the correct order in the attribute of entities. To preserve order, Trisedya et al. have published three papers on content planning from general to specific and thoroughly demonstrated their attention-based solution. The first paper proposed an end-to-end NLG model that implemented a content-plan-based bag of tokens attention model to efficiently capture the order of salient attributes [9]. The second paper focuses more on the encoder part of the model, the GCP graphic encoder with content planning was proposed to preserve the information in the graph when inputting into the model. The content planner was integrated into the encoder by a proposed entity-order aware topological traversal algorithm [10]. The third paper proposed a novel GSC-attention model that captures the local and global context of input attribute to mask only the most important information to be generated in the correct order in the output natural sentence [11].

In this literature review, we have followed Trisedya and his team's footsteps starting by proposing a new approach and further building on it step by step to improve the model's performance. Subsequent to proposing the GTR-LSTM encoder-decoder-based model, Trisedya and his team have worked on entity extraction, entity alignment with a benchmark, and content planning. After familiarizing with the research field and its current progress, we propose to make our contribution to improving the NLG solution by utilizing the pre-trained model technology that has been much-accounted for recently in the NLP field.

## Bibliography


[1] S. Mystakidis, “Metaverse,” *Encyclopedia*, vol. 2, no. 1, Art. no. 1, Mar. 2022, doi: 10.3390/encyclopedia2010031.

[2] “Metaverse Market Size by Technology, Component, End-User, Regions, Global Industry Analysis, Share, Growth, Trends, and Forecast 2022 to 2030 | The Brainy Insights.” https://www.thebrainyinsights.com/report/metaverse-market-12815 (accessed Sep. 04, 2022).

[3] P. Cimiano, J. Lüker, D. Nagel, and C. Unger, “Exploiting Ontology Lexica for Generating Natural Language Texts from RDF Data,” in *Proceedings of the 14th European Workshop on Natural Language Generation*, Sofia, Bulgaria, 2013, pp. 10–19. Accessed: Sep. 05, 2022. [Online]. Available: https://aclanthology.org/W13-2102

[4] B. D. Trisedya, J. Qi, R. Zhang, and W. Wang, “GTR-LSTM: A Triple Encoder for Sentence Generation from RDF Data,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Melbourne, Australia, 2018, pp. 1627–1637. doi: 10.18653/v1/P18-1151.

[5] B. D. Trisedya, G. Weikum, J. Qi, and R. Zhang, “Neural Relation Extraction for Knowledge Base Enrichment,” in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy, 2019, pp. 229–240. doi: 10.18653/v1/P19-1023.

[6] B. D. Trisedya, J. Qi, and R. Zhang, “Entity Alignment between Knowledge Graphs Using Attribute Embeddings,” *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 01, Art. no. 01, Jul. 2019, doi: 10.1609/aaai.v33i01.3301297.

[7] R. Zhang, B. D. Trisedy, M. Li, Y. Jiang, and J. Qi, “A Benchmark and Comprehensive Survey on Knowledge Graph Entity Alignment via Representation Learning.” arXiv, May 05, 2022. doi: 10.48550/arXiv.2103.15059.

[8] A. Vaswani *et al.*, “Attention is All you Need,” in *Advances in Neural Information Processing Systems*, 2017, vol. 30. Accessed: Sep. 04, 2022. [Online]. Available: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

[9] B. Trisedya, J. Qi, and R. Zhang, “Sentence Generation for Entity Description with Content-Plan Attention,” *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 05, Art. no. 05, Apr. 2020, doi: 10.1609/aaai.v34i05.6439.

[10] B. D. Trisedya, J. Qi, W. Wang, and R. Zhang, "GCP: Graph Encoder with Content-Planning for Sentence Generation from Knowledge Base," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–1, 2021, doi: 10.1109/TPAMI.2021.3118703.

[11] B. D. Trisedya, X. Wang, J. Qi, R. Zhang, and Q. Cui, "Grouped-Attention for Content-Selection and Content-Plan Generation," in *Findings of the Association for Computational Linguistics: EMNLP 2021*, Punta Cana, Dominican Republic, 2021, pp. 1935–1944. doi: 10.18653/v1/2021.findings-emnlp.166.

[12] D. Vrandečić and M. Krötzsch, "Wikidata: a free collaborative knowledgebase," *Commun. ACM*, vol. 57, no. 10, pp. 78–85, Sep. 2014, doi: 10.1145/2629489.

[13] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, "Freebase: a collaboratively created graph database for structuring human knowledge," in *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, New York, NY, USA, 2008, pp. 1247–1250. doi: 10.1145/1376616.1376746.

[14] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, 2014, pp. 1724–1734. doi: 10.3115/v1/D14-1179.

[15] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a Method for Automatic Evaluation of Machine Translation," in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, Philadelphia, Pennsylvania, USA, 2002, pp. 311–318. doi: 10.3115/1073083.1073135.

[16] M. Denkowski and A. Lavie, "Meteor 1.3: Automatic Metric for Reliable Optimization and Evaluation of Machine Translation Systems," in *Proceedings of the Sixth Workshop on Statistical Machine Translation*, Edinburgh, Scotland, 2011, pp. 85–91. Accessed: Sep. 25, 2022. [Online]. Available: https://aclanthology.org/W11-2107

[17] M. Snover, B. Dorr, R. Schwartz, L. Micciulla, and J. Makhoul, "A Study of Translation Edit Rate with Targeted Human Annotation," in *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers*, Cambridge, Massachusetts, USA, Aug. 2006, pp. 223–231. Accessed: Sep. 25, 2022. [Online]. Available: https://aclanthology.org/2006.amta-papers.25

[18] A.-C. N. Ngomo and S. Auer, "LIMES - A Time-Efficient Approach for Large-Scale Link Discovery on the Web of Data," p. 6.

[19] M. Chen, Y. Tian, M. Yang, and C. Zaniolo, "Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment," in *Proceedings of the Twenty-*

*Sixth International Joint Conference on Artificial Intelligence*, Melbourne, Australia, Aug. 2017, pp. 1511–1517. doi: 10.24963/ijcai.2017/209.