Jeong-Jin Kang
Edward J. Rothwell
Yang Hao
Sang-Hyun Lee

Advanced and Applied Convergence Letters AACL 22

Advanced and Applied Convergence & Adavanced Culture Technology

11th International Symposium, ISAAC 2023
in Conjunction with ICACT 2023, ICKAI 2023

November 16 - 18, 2023, Jeju, Korea
Revised Selected Papers

IIBC The Institute of Internet, Broadcasting and Communication
IPACT The International Promotion Agency of Culture Technology
지식의 숲 FOREST OF KNOWLEDGE

Extracted Practices on Identifying a Structured Object Diagram with Natural Language Story Scripts based on Linguistic Analysis Mechanism

Janghwan Kim[*], Byung Kook Jeon[**], R. Young Chul Kim[***]

[*,***]*Department of Software and Communications Engineering, Hongik University*
[**]*Department of Software, Gangneung-Wonju National University*
janghwan.kim@g.hongik.ac.kr, jeonbk@gwnu.ac.kr, bob@hongik.ac.kr

Abstract

Recent studies focusing on converting natural language-based text into visual formats have gained attention, particularly the representation of narratives and characters from literary works in formats such as comics or films, which hold significant importance. However, the challenge of visualizing the sophisticated meanings and expressions of text using generative artificial intelligence remains unresolved. To address this issue, we propose a new methodology that reduces the ambiguity of prompts and enhances the transition from text to visual representation. Our methodology integrates linguistic analysis and software engineering design techniques to define a process that accurately analyzes complex text information and connects it to visual creation. This research is expected to contribute to more effectively converting natural language-based novels or texts into visual forms.

Keywords: *Natural Language Understanding, Linguistical Image Design, Object Diagram, Fillmore Semantic Role*.

1. Introduction

The rapid advancement of artificial intelligence (AI) technology is being applied across various fields by researchers worldwide. In particular, research on generative artificial intelligence (Generative AI) has garnered much interest due to its capability to autonomously generate text, images, signals, and sound [1]. Such AI systems generate outputs based on user inputs known as prompts. Most prompts consist of sentences or paragraphs, and users cannot predict the outcomes these prompts will yield [2]. This unpredictability is attributed to the inherent uniqueness of natural language used as prompts [3].

In this paper, we propose a novel approach that reduces the ambiguity of prompts and integrates linguistic analysis with software engineering design techniques for converting text to visual imagery. Through this method, we aim to effectively analyze complex text information and propose a way to generate visual images based on this analysis. This study is expected to contribute to research on effectively transforming natural language-based novels and other texts into visual forms.

In Chapter 2, we discuss related works concerning the linguistic analysis mechanisms of text. Chapter 3 introduces a design mapping mechanism that converts text information into object diagrams through linguistic analysis of natural language and bases image generation on this. Finally, Chapter 4 addresses the conclusions of our research and directions for future research.

2. Related Works

Research on the linguistic analysis of sentences prominently features the work of Noam Chomsky and Charles J. Fillmore. Chomsky, often called the father of modern linguistics, laid the groundwork for a scientific approach to language meaning through structural analysis [4]. His research focused on uncovering the inherent rules and patterns of language, significantly contributing to the structural interpretation of text [5]. On the other hand, Charles Fillmore proposed Case Grammar to explain the meanings of elements within sentences [6]. Fillmore's Case Grammar is an approach to sentence interpretation that centers around semantic roles rather

J.J. Kang et al. (Eds.): ISAAC 2023, AACL 22, 2023

than structural perspectives [7]. This helps in understanding the meaning of a sentence by analyzing the semantic relationships between nouns around the verb. Figure 1 illustrates examples of sentence analysis applying the methods of both linguists.

3. Paper title and author information

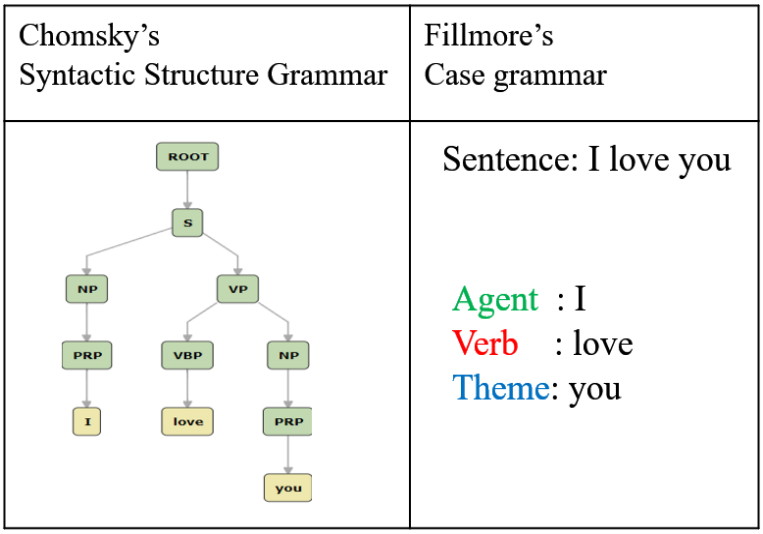

Figure 1. Examples of linguistic analysis methods by Chomsky and Fillmore

3.1. Visualization element mapping mechanism through natural language-based text linguistic analysis

Integrating linguistic analysis and software engineering design techniques in the transition from text to visual content works through the following process. First, the text is analyzed using linguistic analysis methods to understand the story's structure.

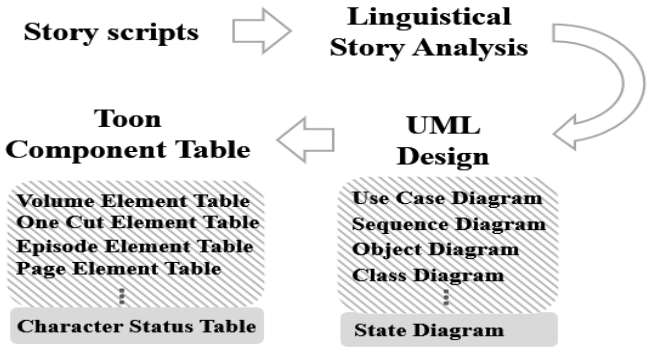

Figure 2. Text information-based visualization element mapping mechanism

3.2. Linguistics-Based Process for Extracting Visualization Elements from Text

In this paper, information about character status changes is extracted from various text script elements. To analyze changes in a character's state by applying software engineering techniques, we use object diagrams to identify objects among UML diagrams.

The following outlines the process of sentence analysis using linguistic analysis:

1. Select the target that will represent a state change.
2. Extract sentences containing the target object from all sentences.
3. Transform any compound or complex sentences extracted in the previous step into simple sentences.
4. Analyze the story script using Chomsky's Phrase Structure Grammar.
5. Based on the analyzed information, determine the roles of sentence elements through Fillmore's semantic role analysis method.
6. Map the elements of the object diagram to the information extracted using Fillmore's semantic analysis method.
7. Create the object diagram based on the mapped information.

4. Applied Practice

Table 1. Example of object identification through Fillmore's semantic analysis

Text script example
...Jany is a nurse at Healthy Town Hospital. At first, she felt overwhelmed and doubted her abilities. Over time, she adapted to the hospital's procedures. Jany learned through her interactions with patients. ...

Table 1 illustrates a segment of the text script example. An object diagram represents the attributes of the objects in a diagrammatic form. Initially, the target object is determined. In this example, Jany, the protagonist, is set as the target object. Subsequently, sentences containing Jany are extracted from all sentences. If the extracted sentences are not simple, they are converted into simple ones. These sentences are then analyzed using Chomsky's Phrase Structure Grammar to identify the parts of speech for each sentence. Following this, the target object, Jany, is classified. At this stage, the roles of the extracted elements are determined using Fillmore's semantic analysis method to extract the role object information. Figure 3 displays the object identification mapping through Fillmore's semantic analysis.

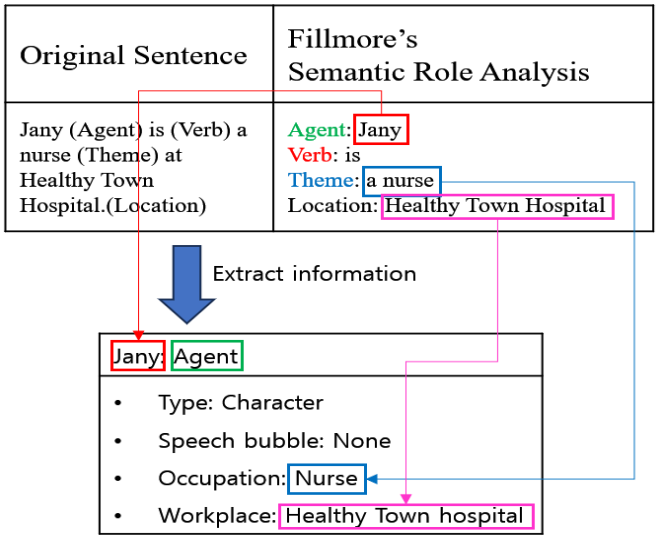

Figure 3. Object diagram mapping mechanism through Fillmore's semantic analysis

This step involves identifying the main verb in the sentence and the roles of the nouns or noun phrases related to the main verb. In the object diagram, information regarding the attributes of the Agent is extracted.

Upon analysis of the sentence, it is revealed that the target object, Jany, is represented with object attribute information such as Type being a character, Occupation as a nurse, and Workplace as Healthy Town Hospital. The object diagram can allow for an intuitive understanding of the object's attributes. Therefore, as the plot unfolds, it becomes possible to comprehend the overall information of the characters present within the text script.

5. Conclusion

In this paper, we propose a methodology that applies linguistic analysis and software engineering design techniques to generate visual content from text. Through this research, we anticipate that the ambiguity inherent in natural language can be improved by applying software engineering design mechanisms.

Future research will apply the methodology of this study to texts of various genres and visual content of different formats to validate their effectiveness and versatility. Additionally, we plan to comprehensively analyze the emotional and psychological impacts of visual content generated by artificial intelligence, aiming to enrich the interaction between humans and AI.

6. Acknowledgement

This research was supported by the Culture, Sports and Tourism R&D Program through the Korea Creative Content Agency (KOCCA) grant funded by the Ministry of Culture, Sports and Tourism (MCST) in 2023(Project Name: Development of AI-based user interactive multi-modal interactive storytelling 3D scene authoring technology, Project Number: RS-2023-00227917, Contribution Rate: 100%)

References

[1] A. Oussidi and A. Elhassouny, "Deep generative models: Survey," 2018 International Conference on Intelligent Systems and Computer Vision (ISCV), Fez, Morocco, 2018, pp. 1-8, doi: 10.1109/ISACV.2018.8354080.B. Sklar, Digital Communications, Prentice Hall, pp. 187, 1998.

[2] H. Dang, L. Mecke, F. Lehmann, S. Goller, and D. Buscheket, "How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models," arXiv preprint arXiv:2209.01390, 2022.

[3] M, Alessio, et al. "What makes my model perplexed? a linguistic investigation on neural language models perplexity." Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures. 2021.

[4] N. Chomsky, "Systems of syntactic analysis." The Journal of Symbolic Logic 18.3 (1953): 242-256.

[5] N. Chomsky, Syntactic structures. Mouton de Gruyter, 2002.

[6] C. J. Fillmore, "The case for case." (1967).

[7] C. J. Fillmore, "Frame semantics." Cognitive linguistics: Basic readings 34 (2006): 373-400.