Classifying Themes of Natural Language-based Sentences based on C3Tree Model for 3D Scene Authoring

Woo Sung Jang
Department of Software and Communications Engineering
Hongik University
Sejong, Korea
uriel200@hongik.ac.kr

Jihoon Kong
Research Institute
Toonsquare
Seoul, Korea
john.tooning@toonsquare.co

Keun-Sang Yi
Jeonbuk Technopark
Jeonju, Korea
yi@jbtp.or.kr

Jang Hwan Kim
Department of Software and Communications Engineering
Hongik University
Sejong, Korea
janghwan.kim@g.hongik.ac.kr

Hyuntae Kim
Department of Software and Communications Engineering
Hongik University
Sejong, Korea
hyuntaekim@g.hongik.ac.kr

Ye Jin Jin
Department of Software and Communications Engineering
Hongik University
Sejong, Korea
yejin_jin@g.hongik.ac.kr

Kidu Kim
Telecommunications Technology Association
Seongnam, Korea
kdkim@tta.or.kr

R. Young Chul Kim
Department of Software and Communications Engineering
Hongik University
Sejong, Korea
bob@hongik.ac.kr

***Abstract*— Currently, we must spend a lot of effort and time to create 3D scenes. Most cartoon writers take a long time to draw objects and backgrounds apart from their particular passions. To solve this problem, we are considering AI-based 3D scene authoring based on multimodal interactive storytelling technology. This research automatically generates 3D scenes with natural language. We, as part of our whole research, propose an AI learning method to identify 3D scene themes from natural language sentences. That is, we make a learning model to generate the C3Tree model from natural language sentences and map terminal nodes (simple sentences) of the C3Tree Model with themes based on Korean Language Understanding Evaluation (KLUE) Topics. This C3Tree model generates simplified sentences with complex sentences. Consequently, we can increase the clarity of data for learning. With this approach, we expect to provide previous datasets of the existing 3D objects and backgrounds to toon-experts for quick reference and also an easy 3D scene creation method for non-experts.**

Keywords—C3TreeModel, 3D Scene, User Interactive Multi-Modal, 3D Scene Authoring

I. Introduction

Recent advancements in 3D production technology and natural language understanding have been extensively researched. For example, The Kaedim makes a service to generate 3D from 2D [1]. The NVIDIA develops GET3D to create 3D model renderings from images[2]. The Sketchfab offers a platform for creating 3D environments and avatars [3]. The KLUE [4] and AI HUB [5] release text and voice data for Korean-based natural language processing technologies.

However, the existing research lacks constructing a 3D authoring environment for webtoon creation. There is no established learning data for storytelling or 3D scene authoring. At this moment, there is also insufficient technology for easy and fast creation of 3D scenes for understanding user intentions.

To solve this problem, we propose AI-based 3D scene authoring multimodal interactive storytelling technology. This method supports experts to quickly create reference 3D objects & scenes, and enables non-experts to easily produce 3D scenes. When a user inputs a natural language sentence into this tool, the meaning of the sentence is analyzed, and suitable 3D scenes and objects are automatically generated based on the sentence's meaning.

This paper, as part of the research on AI-based 3D scene authoring multimodal interactive storytelling technology, proposes a method for selecting themes for 3D scenes using AI learning of natural language sentence themes. This method uses simplified natural language sentences with Conditional and Conjunction Clause Tree (C3Tree) modeling [6] to learn 3D scene themes. We also calculate the F1 Score based on prediction results. This can increase the accuracy of 3D scene authoring using natural language sentences and quantify the accuracy of the model.

This paper is as follows: Chapter 2 mentions related research. Chapter 3 discusses our research. Chapter 4 talks about experiments and results. Finally, the conclusion is mentioned.

II. Related Studies

A. C3Tree Model

The C3Tree Model represents the process of simplifying complex sentences into simple sentences using the tree. It separates complex sentences into clauses based on sentence normalization. During the clause separation process, the omitted subjects are restored. Passive sentences are converted

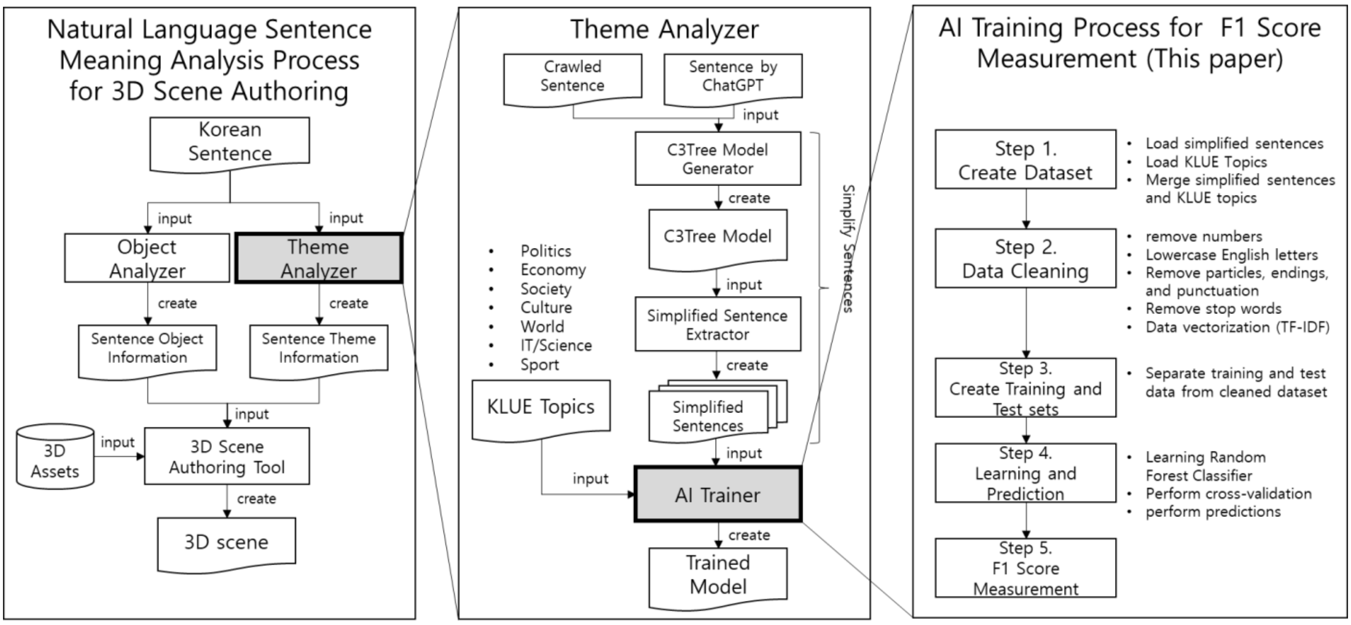

Fig. 1. Natural Language Sentence Meaning Analysis Process for 3D Scene Authoring

into active ones. Simplified sentences with similar meanings are integrated into one.

Figure 1 is an example of the C3Tree Model. The two top nodes are the original natural language sentences. The middle nodes are compound sentences that are not fully simplified. The terminal nodes are fully simplified sentences. The relationships between each node are constituted as either a positive cause/effect relationship, a negative cause/effect relationship, an AND relationship, or an OR relationship.

This model provides traceability in the sentence simplification process.

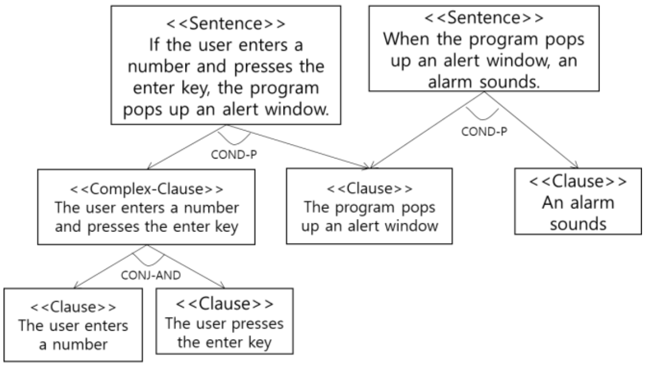

Fig. 1. Example of C3Tree Model

B. KLUE Benchmark

KLUE[4] is a benchmark suite for evaluating Korean natural language understanding. It provides high-quality evaluation datasets and automated metrics for system performance testing. The table I shows the measurement items included in KLUE.

TABLE I. PART OF TRAINING SET

Name	Description
Topic Classification	This provide a classifier for predicting the topic of text snippets.
Sementic Textual Similarity	This is measure the degree of semantic equivalence between two sentence.
Natural Language Inference	This reads pairs of whole sentences and hypothesis sentences, and predicts whether the relationship is contradictory or neutral.
Named Entity Recognition	This is detect the boundaries of named entities in unstructured text and classify the types.
Relation Extraction	This identifies semantic relations between entity pairs in a text.
Dependency Parsing	This finding relational information among words
Machine Reading Comprehension	This evaluates the ability to understand questions and find answers.
Dialogue State Tracking	This is about predicting the dialogue states from a given dialogue context.

III. CLASSIFYING THEMES OF SENTENCES BASED ON THE C3TREE MODEL FOR 3D SCENE AUTHORING

Figure 2 shows the full research process. The left process in Fig 2 identifies object and theme information from natural language-based Korean sentences and uses this to create a 3D scene. The middle one in Figure 2 is the process of the theme analyzer, which simplifies the original text with the C3Tree Model. We learn the AI model with simplified sentences generated by the C3Tree model for AI learning. In this paper, we just mention the AI training process in the right one of Figure 2. The detailed steps are as follows:

Step 1. Map the simplified Korean sentences into one of the appropriate KLUE Topics. KLUE's Topic Classification provides a benchmark standard for Korean language topic classification. The types of KLUE Topics include Politics, Economy, Society, Culture, World, IT/Science, and Sports. This step creates a large dataset.

Table II is the index number for each topic. A unique number is assigned to 7 topics.

TABLE II. INDEX OF TOPIC

Index	Topic
0	Politics
1	Economy
2	Society
3	Culture
4	World
5	IT/Science
6	Sports

Table III is part of the mapped data set. It consists of a sentence index, Korean natural language sentence, and topic index.

TABLE III. PART OF MAPPED DATA SET

Index	Korean Sentence (Pronunciation)	Topic Index
0	도시의 중심가에서 사람들이 경제적 불평등에 대해 시위하다. (dosiui jungsimga-eseo salamdeul-i gyeongjejeog bulpyeongdeung-e daehae siwihada.)	0
1	어른들이 공공 도서관에서 사회 변화를 주제로 한 책을 읽고 토론하다. (eoleundeul-i gong-gong doseogwan-eseo sahoe byeonhwaleul jujelo han chaeg-eul ilg-go tolonhada.)	0

Step 2. Clean the dataset. Remove numbers. Lowercase English. Eliminate particles, endings, punctuation. Remove 17 stopwords. Use TF-IDF [7] to vectorize the data.

Step 3. Separate the clean dataset into a training set and a test set.

Step 4. Run learning based on a Random Forest Classifier [8]. Perform cross-validation and predict using the test set.

Step 5. Calculate the F1 Score based on the prediction results.

IV. EXPERIMENT AND RESULTS

We train the proposed model with Korean sentences and evaluate the results. The sentences in Table IV are part of the training sentences.

TABLE IV. PART OF SENTENCE IN TRAINING SET

Korean (Pronunciation)	English
도시의 중심가에서 사람들이 경제적 불평등에 대해 시위하다. (dosiui jungsimga-eseo salamdeul-i gyeongjejeog bulpyeongdeung-e daehae siwihada.)	People protest against economic inequality in city centers.
어른들이 공공 도서관에서 사회 변화를 주제로 한 책을 읽고 토론하다. (eoleundeul-i gong-gong doseogwan-eseo sahoe byeonhwaleul jujelo han chaeg-eul ilg-go tolonhada.)	Adults read and discuss books on social change at a public library.
친구들이 쇼핑몰의 푸드코트에서 다양한 음식을 먹으며 이야기를 나누다. (chingudeul-i syopingmol-ui pudeukoteueseo dayanghan eumsig-eul meog-eumyeo iyagileul nanuda.)	Friends chatting while eating various foods at a shopping mall's food court.

Table V shows the number of datasets. The training set is 1680, and the test set is 280.

TABLE V. COUNT OF DATA SET

Name	Count
Training set	1680
Test set	280

Table VI shows the number of training data per topic. We input 240 sentences evenly for each topic.

TABLE VI. NUMBER OF TRAINING DATA PER TOPIC

Topic	Count
Politics	240
Economy	240
Society	240
Culture	240
World	240
IT/Science	240
Sports	240

Table VII is the prediction success rate per topic.

TABLE VII. PREDICTION SUCCESS RATE PER TOPIC

Topic	Prediction Rate
Politics	100%
Economy	100%
Society	100%
Culture	100%
World	100%
IT/Science	100%
Sports	100%

Equation (1) is the prediction result of the learned model. All sentences were predicted correctly.

$$\text{Prediction Rate} = 1.0 \tag{1}$$

Equation (2) calculates Precision using the prediction rate. Precision is the probability that the model predicts as True is actually True.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} = \frac{1.0}{1.0+0.0} = 1.0 \tag{2}$$

Equation (3) calculates Recall using the prediction rate. Recall is the probability that the model correctly identifies actual True data as True.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negitives}} = \frac{1.0}{1.0+0.0} = 1.0 \tag{3}$$

Equation (4) is the F1 score calculated using precision and recall. The F1 score is the harmonic mean of precision and recall.

$$\text{F1 Score} = 2 * \frac{\text{Precision}*\text{Recall}}{\text{Precision+Recall}} = 2 * \frac{1}{2} = 1.0 \tag{4}$$

V. CONCLUSIONS

We propose a method for selecting themes for 3D scenes and AI learning of natural language sentence themes, which is a part of our research on AI-based 3D scene authoring multimodal interactive storytelling technology. This method normalizes and simplifies sentences, and trains them after TF-IDF vectorization. We ultimately enhances the clarity of the training data.

As a result of our experiment, the prediction rate is 100%. Therefore, the F1 Score was calculated as 1.0.

For future research, we plan to collect, create, evaluate more datasets for theme analysis, and also make a method for object analysis from natural language sentences.

ACKNOWLEDGMENT

This research was supported by 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: 50%) and the BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education(MOE, Korea) and National Research Foundation of Korea(NRF) in 2023(Project Name: Research Team for Super-Distributed Autonomous Computing Service Technologies, Project Number: 202003520004, Contribution Rate: 50%).

REFERENCES

[1] KAEDIM, https://www.kaedim3d.com (accessed on 15 November 2023).

[2] NVIDIA GET3D, https://research.nvidia.com/labs/toronto-ai/GET3D (accessed on 15 November 2023).

[3] Sketchfab, https://sketchfab.com (accessed on 15 November 2023).

[4] KLUE, https://klue-benchmark.com (accessed on 15 November 2023).

[5] AI Hub, https://www.aihub.or.kr (accessed on 15 November 2023).

[6] W. S. Jang, and R. Y. C. Kim, "Automatic Generation Mechanism of Cause-Effect Graph with Informal Requirement Specification Based on the Korean Language", Applied Sciences, Vol.11, No.24, pp. 1-13, December 2021.

[7] A. Aizawa, "An information-theoretic perspective of tf–idf measures", Information Processing and Management, Vol.39, Issue.1, pp.45-65, September 2002.

[8] L. Breiman, "Random Forests", Machine Learning, Vol.45, pp. 5–32, October 2001.