The Transactions of the Korea Information Processing Society
TKIPS Vol.13, No.10, pp.529~536

eISSN: 3022-7011
https://doi.org/10.3745/TKIPS.2024.13.10.529

3D Object State Extraction Through Adjective Analysis from Informal Requirements Specs

Ye Jin Jin† · Chae Yun Seo†† · Ji Hoon Kong††† · R. Young Chul Kim††††

ABSTRACT

Recent advancements in AI technology have led to its application across various fields. However, the lack of transparency in AI operations makes it challenging to guarantee the quality of its outputs. Therefore, we integrate requirements engineering in software engineering with conversational AI technology to ensure procedural fairness. Traditional requirements engineering research uses grammar-centered analysis, which often fails to fully interpret the semantic aspects of natural language. To solve this, we suggest combining Noam Chomsky's syntactic structure analysis with Charles Fillmore's semantic role theory. Additionally, we extend our previous research by analyzing adjectives in informal requirement sentence structures. This enables precise emotional analysis of the main characters in comics. Based on the results of the analysis, we apply the emotional states of the objects to the states in the UML state diagram. Then, we create the 3D object with Three.js based on the object that reflects the emotional states in the state diagram. With this approach, we expect to represent the emotional state of a 3D object.

Keywords : Informal Requirements, UML State Diagram, 3D Object Model, Adjective Analysis

비정형 요구사항 스펙에서 형용사 분석을 통한 3D 객체 상태 추출화

진 예 진† · 서 채 연†† · 공 지 훈††† · 김 영 철††††

요 약

최근 AI 기술의 발전으로 다양한 분야에서 활용되고 있다. 그러나 AI의 동작 원리를 확인할 수 없어 생성물의 품질을 보장하기 어렵다. 따라서 본 연구에서는 대화형 AI 기술에 소프트웨어 공학의 요구공학을 접목하여 절차적 공정을 진행한다. 기존의 요구공학 연구는 문법 중심 분석 방식을 사용하여 비정형 요구사항의 의미적 측면을 충분히 해석하지 못하는 한계를 가진다. 이를 해결하기 위해, 우리는 촘스키의 구문 구조 분석 이론과 필모어의 의미역 이론의 통합을 제안한다. 또한, 우리는 이전 연구를 확장하여 다양한 문장 구조에서 사용되는 형용사 분석을 추가한다. 이는 만화 속 주인공에 대한 정확한 감정 분석이 가능하다. 분석 결과를 기반으로, 우리는 상태 다이어그램의 상태를 객체의 감정 상태로 적용한다. 우리는 객체의 상태 다이어그램의 감정 상태 기반으로 Three.js를 통해, 3D 객체 모델 생성과 와 상태를 표현시킨다. 이를 통해, 객체의 감정을 표현하기를 기대한다.

키워드 : 비정형 요구사항, 상태 다이어그램, 3D 모델, 형용사 분석

※ 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 multimodal interactive storytelling 3D scene authoring technology, Project Number: RS-2023-00227917, Contribution Rate: 80%) and Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education in 2023(Project Name: The Faultless Codinglization through Automatic Refactoring based on NLP BERT Model, Project Number: 2021R1I1A3050407, Contribution Rate: 10%) and BK21 FOUR (Fostering Outstanding Universities for Research) funded by the Ministry of Education (MOE, Korea) in 2023 (Project Number: 2021R1I1A305040711, Contribution Rate: 10%).

※ 이 논문은 2024년 ASK 2024의 일반논문으로 "비정형 자연어 요구 사항 기반 상태 모델을 통한 3D 객체의 상태 표현 메커니즘"의 제목으로 발표된 논문을 확장한 것임.

† 준 회 원 : 홍익대학교 소프트웨어공학연구실 석사과정
†† 비 회 원 : 홍익대학교 소프트웨어융합학과 초빙교수
††† 비 회 원 : 홍익대학교 소프트웨어공학연구실 박사과정
†††† 정 회 원 : 홍익대학교 소프트웨어융합학과 정교수

Manuscript Received : July 31, 2024
First Revision : September 20, 2024
Accepted : September 23, 2024

* Corresponding Author : R. Young Chul Kim(bob@hongik.ac.kr)

1. Introduction

With the advancement of AI technology in AI industrial fields, traditional tasks performed manually in various fields are now being rapidly and easily processed using AI. This can get the desired results without complex technology or specialized knowledge to make the process efficient and productive. Recently, with the rise of AI technologies such as Text to Image (TTI) and Text to Speech (TTS), AI tools of cartoon development industry have also been working in the field of 3D modeling, leading to the development of Text to 3D technology.

Traditionally, 3D modeling demands significant time and advanced technical expertise, but the Text to 3D (TT3D) AI enables the rapid generation of 3D models through natural language without the need for complex modeling skills. However, because generative AI operates as a black box, it is hard to ensure the consistency and quality of its outputs [1]. From a software engineering perspective, this challenges for consistency and reusability with maintenance and quality assurance. To resolve these issues, previous research has proposed mechanisms for generating 3D objects from informal requirements in comics or webtoons [2]. The previous research integrates requirements engineering and linguistic theories to analyze the semantics of natural language requirements effectively. Based on the textual analysis, we generated state diagrams to represent the character's emotional change and then created corresponding objects reflecting these emotions. However, these methods primarily focused on verbs to analyze emotions, thus failing to capture the detailed nuances of emotions. We did not consider the role of adjectives, leading to a lack of precision in emotion analysis. This study aims to overcome these limitations by extending previous research. By integrating Chomsky's syntactic structure analysis with Fillmore's case grammar theory and also including adjective analysis, we aim to capture the detailed nuances of emotions more accurately in natural language requirement specifications. We propose a method to reflect the emotional states in 3D objects more precisely, ensuring a higher degree of reusability, accuracy, and consistency in the generated outputs.

The structure of this paper is as follows: Section 2 discusses the relevant linguistics and adjective analysis, as well as existing research. Section 3 describes the mechanisms proposed in this study. Section 4 compares our approach with other text-to-3D AI tools. Section 5 concludes with the expected outcomes and implications of the research and mentions future research.

2. Related Research

2.1 Our Previous Approach

In our previous study, character emotions were extracted through emotion verbs in natural language. We represented emotions in state diagrams to visualize emotional transitions effectively [2]. Based on this method, we created 3D objects that displayed changes in facial expressions. Fig. 1 illustrates the mechanism proposed in our previous study.

The previous method categorized adjectives following stative verbs as indicators of emotion within the sentences. However, this approach only pays attention to adjectives following state verbs. Adjectives have other structures as well. In particular, adjectives that directly modify nouns cannot be interpreted in the previous study. To address this gap, we focus on a systematic analysis of adjectives, enabling more precise emotion extraction. This enhancement aims to refine the accuracy and comprehensiveness of our emotional analysis, thereby improving the emotional representation in both state diagrams and 3D models.

2.2 The Linguistic Theories Utilized in This Study

Previous study have used two different linguistics, Chomsky and Fillmore, to analyze English sentences for requirement engineering. The same linguistics are applied in this study. The Chomsky approach focuses on syntactic structure analysis, but the Fillmore approach focuses on

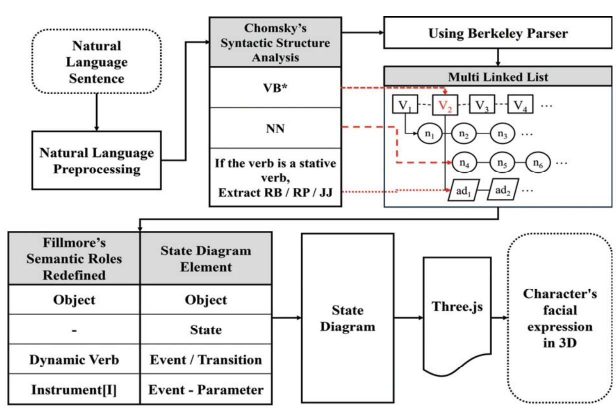
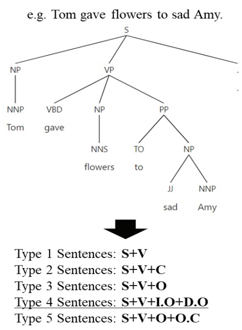
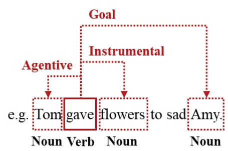


Fig. 1. Mechanism of Generating 3D Pictures from Natural Language in Existing Research

Table 1. Comparison with Two Linguistic Approaches

Criteria	Chomsky's Syntactic Structure Analysis	Fillmore's Case Grammar
Key Features	The sentence structure and part of speech can be identified.	Nouns associated with the main verb are assigned semantic roles.
Limitations	Semantic interpretation of the sentence is not feasible.	Structural interpretation of the sentence is not feasible.
Illustrative Example		

case grammar for semantic analysis [3]. Table 1 represents two linguistic approaches.

1) Chomsky's Syntactic Structure Analysis

Noam Chomsky's theory of syntactic structures aims to understand the architecture of language and systematically analyze grammatical constructions [4]. We use the Berkeley Parser tool based on Chomsky's theory to analyze natural language sentences. This parser transforms natural language sentences into tree structures for textual analysis [5]. These tree structures enable the identification of sentence components through part-of-speech tags. Table 2 depicts commonly used part-of-speech tags in our research.

Table 2. Part-of-speech Tags

Notation	Tag	Description
NN*	NN	Noun, singular or mass
	NNS	Noun, plural
	NNP	Proper noun, singular
	NNPS	Proper noun, plural
VB*	VB	Verb, base form
	VBD	Verb, past tense
	VBG	Verb, gerund or present participle
	VBN	Verb, past participle
	VBP	Verb, non-3rd person singular presnet
	VBZ	Verb, 3rd person singular present
JJ*	JJ	Adjective
	JJR	Adjective, comparative
	JJS	Adjective, superlative

Table 3. Part of the Original Fillmore's Case Grammar

Case	Description
Agentive	The instigator of the event
Objective	This refers to the things affected by the state or the thing determined by the verb.
Instrumental	This refers to an inanimate force or object involved as a factor for the action or state determined by the verb.
Dative	This refers to a living being affected by an action or state determined by the verb.
Benefactive	This refers to the living object for which the action determined by the verb serves.

The asterisk(*) in Table 2 represents the consolidation of various derived part-of-speech tags into their basic categories. For example, 'VBG' indicates the present progressive form of a verb, and 'VBD' represents the past tense. These are considered derivatives of the basic verb form 'VB' and are denoted as 'VB*.' In our study, this notation is employed to analyze sentences efficiently.

2) Fillmore's Case Grammar

Charles J. Fillmore's case grammar theory aims to identify semantic relationships between nouns and a main verb within sentences [6]. This explains how sentence components interact based on the semantic roles of a main verb, allowing for a deeper analysis of syntactic meaning. Table 3 provides some of the role descriptions defined in the Fillmore case grammar.

He offers a broad and diverse approach to interpreting meaning, allowing for versatile applications. Therefore, we refine his theory to utilize it appropriately within our mechanisms.

2.3 Adjective Analysis

We identify the adjective structure to analyze the emotions that appear in various sentences. Previous studies have identified only state verbs and adjectives following them for sentiment analysis. However, adjectives have many different structures. Adjectives represent structural characteristics within sentences that differ from nouns and verbs [7]. Adjectives that function as direct attribute modifiers of nouns directly modify the nouns and express their characteristics and states. This structure is unique to adjectives and cannot be found in nouns or verbs. Using these features, we identify adjective structures and perform sentiment analysis more effectively than before.

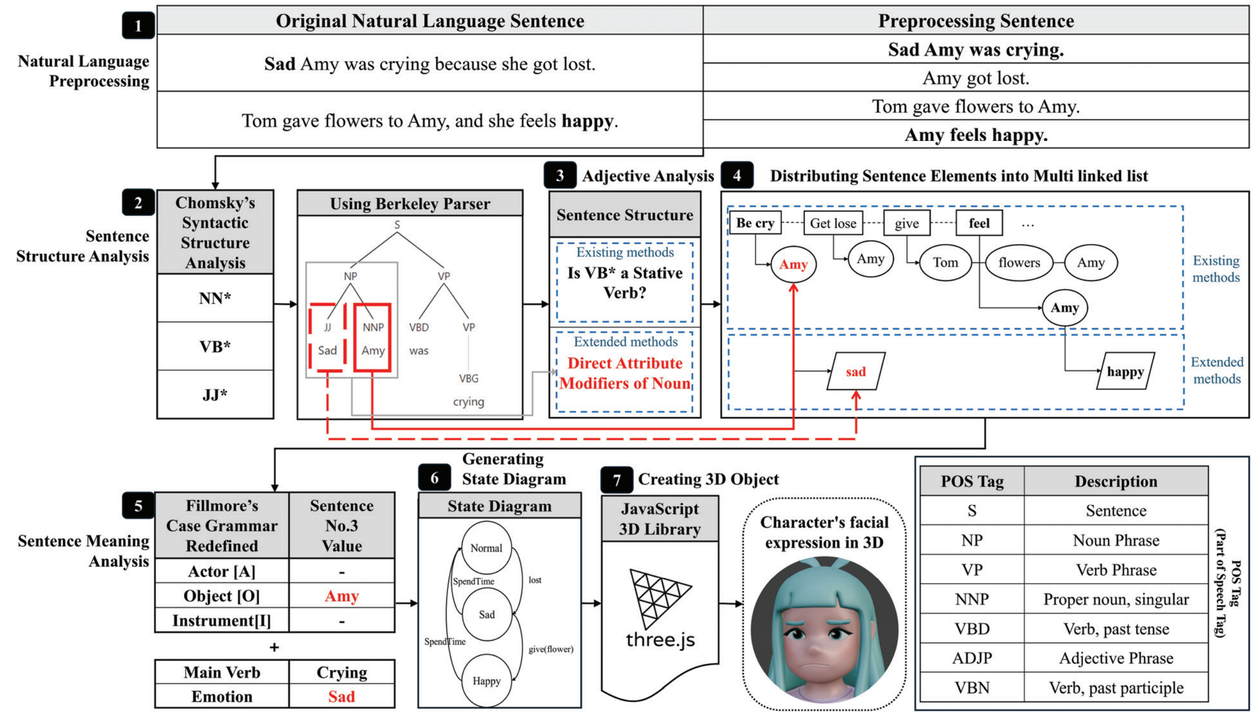


Fig. 2. Overview of Our Whole Cartoon Generation Process

3. 3D Object Generation via State Diagram Through Informal Requirement Specification Analysis

We propose a mechanism for generating 3D models via state diagrams through natural language processing. The process comprises systematic natural language analysis, the creation of state diagrams, and the generation of 3D objects reflecting emotional states. In particular, we extend the adjective analysis structure of existing studies to enable analysis of more emotions. Fig. 2 illustrates our overall process.

1) Natural Language Preprocessing

At number 1 of Fig. 2, ensuring accurate input natural language text analysis requires a preprocessing step. Compound or complex sentences complicate the analysis [8]. To improve the accuracy of the analysis, a simplification process is undertaken, transforming complex sentences into simpler ones. In this process, if the subject is omitted, it is reintroduced from the original sentence to form a complete sentence.

2) Chomsky's Syntactic Structure Analysis

At number 2 of Figure 2, the preprocessed sentences are analyzed using the Berkeley Parser, which is based on Noam Chomsky's theory. This parser represents sentences in a tree structure, allowing for the identification of the key verbs, adjectives, and nouns within the sentences.

3) Adjective Analysis

At number 3 of Fig. 2, the analysis of adjectives using the Berkeley Parser covers a specific structure. According to previous studies, emotion analysis is only feasible when using stative verbs followed by emotional adjectives. Therefore, only sentences like 'Amy feels happy.' in Fig. 3 can be analyzed. However, adjectives have various structures, and previous studies did not take this into account. Previously, we could not analyze sentences such as 'Sad Amy was crying' in Fig. 3. However, through the extended analysis in this study, it is also possible to analyze adjectives that directly modify nouns. The structure where adjectives directly modify nouns appears within a noun phrase (NP) containing an adjective (JJ) and a noun (NN).

4) Extended Storage Structure of the Multi-liked list

At number 4 of Fig. 2, the results of the syntactic analysis are stored in a multi-linked list. The list is structured around the sentence's main verb, connecting related

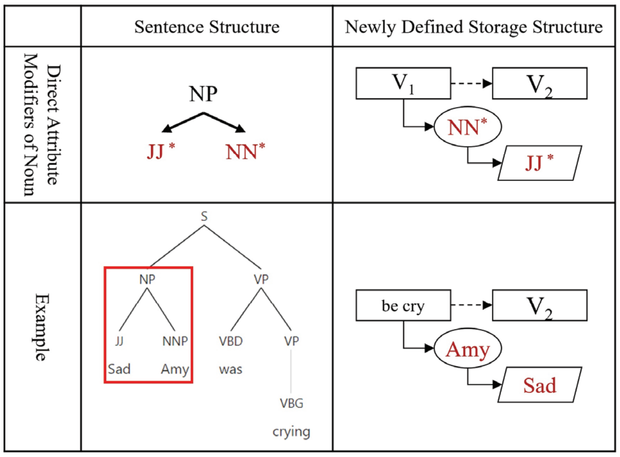


Fig. 3. Storage Structure According to Sentence Tree Structure

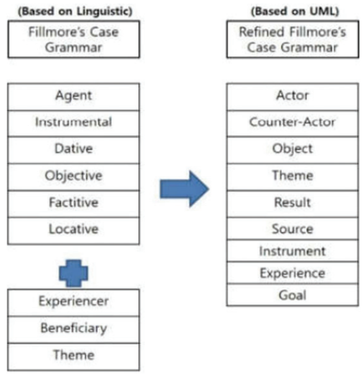
nouns to this verb. Additionally, if there are adjectives modifying these nouns, they are also linked to the respective nouns. In order to confirm these relationships, we analyze sentence structure using a tree structure. Fig. 3 shows how the analyzed verbs, nouns, and adjectives are stored.

The relationship between JJ* and NN*, highlighted in red in Fig. 3, represents the extended part of previous research. The structure of the multi-linked list allows us to identify the modification relationship between adjectives and nouns. An example sentence is provided at the bottom of Fig. 3 for explanation.

5) Fillmore's Case Grammar

At number 5 of Fig. 2, Fillmore's Case grammar theory is applied to the data stored in the multi-linked lists. This analysis allows each element to be examined according to specific case grammar, clearly understanding the sentence's meaning. We apply Fillmore's theory for the generation of state diagrams. However, we follow the redefined Case Grammar from a UML perspective, which differs from Fillmore's original definitions [9]. This study defined the cases required for generating UseCases in UML. Fig. 4 illustrates the redefined case grammar from a UML perspective.

From the nine cases defined in this study, we will use only three cases necessary for generating the State Diagram. The Actor, Object, and Instrument indicated in Fig. 5 are the essential cases for generating the State Diagram. This redefinition helps interpret each sentence component's roles more clearly and generate state diagrams effectively.



Case	Notation	Definition
Actor	A	The instigator of the event/action.
Counter actor	CA	The force or resistance against which an action is carried out.
Object	O	The entity which moves or changes or whose existence is in consideration.
Theme	TO	The subjective entity of Objects
Result	R	"Entity" that comes into existence as a result of the action.
Source	S	Origin of object.
Instrument	I	Facility used in carrying out an event.
Experience	E	"Entity" that receives or accepts or experiences or undergoes the effect of an action.
Goal	G	Destination of object.
Main Verb	V	A change defined by an event.

Fig. 4. Case Grammar Redefined Based on UML[9]

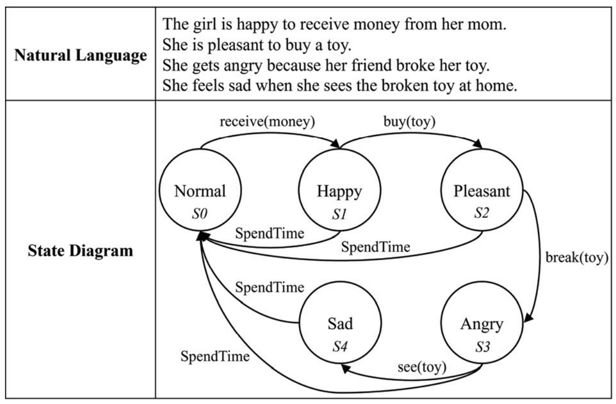


Fig. 5. State Diagram Generation from Natural Language

6) State Diagram Generation

After the natural language analysis is complete at number 6 of Fig. 2, the analyzed sentence elements are mapped to state diagram elements. The objects in the state diagram represent the main character of sentences, and the states represent their emotional states. We complete the state diagram by representing states with circles and events with arrows. Fig. 5 shows how a state diagram is constructed from natural language sentences.

The states defined in our study consist of five categories: Normal, Happy, Pleasant, Sad, and Angry. The state begins at S0, which is Normal, and transitions according to the events triggered by the sequence of sentences. We defined these states as emotions. Since emotions can re-

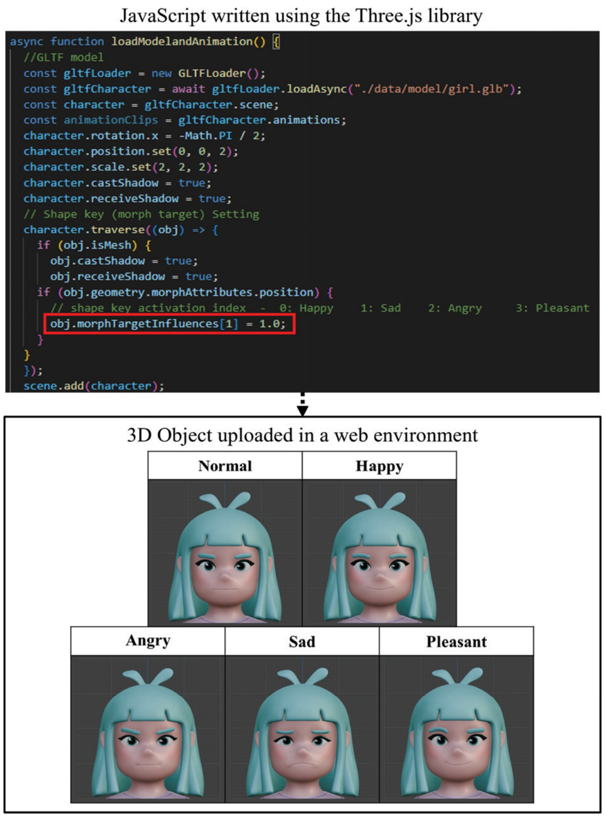


Fig. 6. Facial Expressions of a 3D Model using Three.js

turn to the Normal state after a certain amount of time, we introduced an event called SpendTime to represent the passage of time.

7) 3D Object Generation

We use Visual Studio Code 1.92.2 as the integrated development environment (IDE) for generating 3D objects, Node.js 20.14.0 for executing JavaScript on the server side, and Three.js 0.167.1 for implementing 3D graphics. Additionally, we use Express 4.19.2 to set up the web server and handle data between the client and server.

At number 7 of Fig. 2, based on the state diagram, the 3D model of the main character is uploaded using Three.js, equipped with various facial expressions to depict various emotions. Each model visually represents the emotional states analyzed in the state diagram. Fig. 6 shows the various facial expressions of a main character, distinctly categorized into four emotions: Happy, Angry, Sad, and Pleasant. This visual representation allows users to understand emotional changes effectively and intuitively visualize complex natural language.

Fig. 6 shows some JavaScript code used to generate the 3D object. Specifically, the part marked with a red box

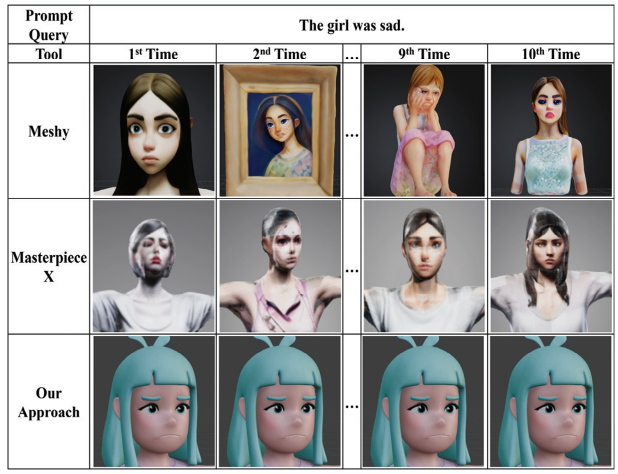


Fig. 7. Comparative Results of Our Approach and the Commercial AI Tools

represents the critical code for emotional expression. The indices of obj.morphTargetInfluences range from 0 to 3, where 0 represents joy, 1 represents sadness, 2 represents anger, and 3 represents pleasure. Additionally, the floating-point value is input into obj.morphTargetInfluences [index] ranges from 0 to 1, where a value closer to 0 reflects a normal expression, and a value closer to 1 represents the emotion at its maximum intensity.

4. Comparison of Result

We compare our mechanism with existing generative AI tools. We established the following criteria to select AI tools appropriate for the research: 1. Can the tool generate 3D models from natural language?; 2. Can the generated 3D models be exported in formats such as .obj, .fbx, or .gltf?; 3. Does the tool work smoothly in a web environment and allow browser-based access and usage?. Among the AI tools that generate 3D models from prompts, Meshy and Masterpiece X meet these criteria. Meshy automatically generates 3D image models when users input text or upload an image [10]. Masterpiece X, built in collaboration with NVIDIA, automates modeling and animation tasks [11]. Fig. 7 presents the comparative results between the proposed mechanism and the AI tools.

We tested inputting the information that a girl character is sad and generating a 3D model of the girl ten times. Our approach consistently produces a 3D model with a sad expression, accurately reflecting the input emotion. In contrast, the generative AI tools often provided models with expressions that did not reflect the specified emo-

Table 4. Comparison of AI Tools and the Results of Our Study

Criteria	Meshy	Masterpiece studio	Our Approach
Can the 3D model express emotions?	△	×	○
Can the 3D model represent various actions?	○	○	○
Can the same 3D model be reused in different contexts?	×	×	○
Can the natural language-based 3D model be validated for consistency through a systematic process?	×	×	○

tion. This discrepancy highlights that despite significant advancements in AI, there still is a gap in understanding and accurately depicting emotions and facial expressions. Furthermore, generative AI, characterized by creativity, produces varied outputs with each attempt. Our research addresses issues of consistency and reusability in character design through a software engineering approach. In cartoon engineering, maintaining consistency in character changes is a critical issue that can be addressed from a software engineering perspective, focusing on reusability and consistency. Our proposed method aligns with these requirements, and Table 4 presents the evaluation results compared to generative AI tools.

As presented in the table, our method delivers effective results in emotional expression, action representation, reusability, and consistency verification with the input meaning. From a comic engineering perspective, maintaining consistency in character expression is essential. Similarly, in software engineering, these factors align with reusability and consistency. Our approach proves effective and superior in ensuring consistent and reusable emotional expressions in comics.

5. Conclusion

The original research focused on automatic code generation through software modeling from English-based sentences. However, in this study, the focus is on generating cartoon images instead of code generation. This is the first attempt to integrate linguistics, requirements engineering, and cartoon engineering. We aim to generate state diagrams through systematic natural language analysis and create 3D object facial expressions to achieve this. By emphasizing the role of adjectives in natural language requirements analysis, we seek to overcome the limitations of traditional verb-centric analysis methods. Based on Chomsky's syntactic structure theory and Fillmore's case grammar theory, we map important attributes of natural language to UML state diagrams to extract the emotional states of objects. These emotions are then visualized using the Three.js library to represent the emotional states of 3D objects. Ultimately, it becomes possible to express the emotional state of an object from natural language. Unlike existing AI tools, this approach addresses issues related to reusability and the development process. In the future, we plan to extend the analysis to handle complex sentence structures, which are difficult even for humans to analyze, and develop this approach as an efficient tool for assisting in cartoon production.

References

[1] J. Gao, C. Tao, D. Jie, and S. Lu, "What is AI software testing? and why," *2019 IEEE International Conference on Service-Oriented System Engineering (SOSE)*, IEEE, pp. 27-36, 2019.

[2] Y. J. Jin, C. Y. Seo, and R. Y. C. Kim, "3D Object State Representation via State Diagram based on Informal Natural Language Requirement Specifications," *Proceedings of the Annual Symposium of Korea Information Processing Society Conference (KIPS) 2024*, Vol.31, pp.494-496, 2024.

[3] Z. Feng, "Formal analysis for natural language processing: a handbook," Springer, pp.520-534, 2023. https://doi.org/10.1007/978-981-16-5172-4.

[4] N. Chomsky, Syntactic structures, USA: Mouton de Gruyter, 2002.

[5] Berkeley Neural Parser [Internet], https://parser.kitaev.io/.

[6] C. J. Fillmore, "The Case for Case," New York: HR&W, 1968.

[7] S. Kasper, "A comparison of 'thematic role' theories," 2008. doi:10.13140/2.1.1540.4805.

[8] B. Das, M. Majumder, and S. Phadikar, "A novel system for generating simple sentences from complex and compound sentences," *International Journal of Modern Education and Computer Science*, Vol.11, No.1, pp.57-64, 2018.

[9] B. K. Park and R. Y. C. Kim, "Effort estimation approach through extracting use cases via informal requirement specifications," *Applied Sciences*, Vol.10, No.9, 2020. doi:10.3390/app10093044

[10] Meshy Documentation [Internet], https://docs.meshy.ai/.

[11] Masterpiece X [Internet], https://www.masterpiecex.com/.

Ye Jin Jin
https://orcid.org/0009-0000-4996-2436
e-mail : yejin_jin@g.hongik.ac.kr

She received a B.S. degree in Software and Communications Engineering from Hongik University in 2023. She has been a master's student in Software Convergence at Hongik University since 2023. Her research interests are in the areas of Software Visualization, Requirement Engineering, Software Engineering, Natural Language Processing, and Artificial Intelligence.

Chae Yun Seo
https://orcid.org/0009-0007-3190-9806
e-mail : chaeyun@hongik.ac.kr

She received the Ph.D. degree in Electronic and Computer Engineering from Hongik University. She worked as an assistant professor in the Department of Internet Information and Communication Engineering at Korea International University in Ferghana. She has been an is currently a research professor at the Metaverse Convergence SW Academy and a visiting Professor in the Department of Software and Communications Engineering at Hongik University. Her research areas include metaverse, AR/VR, testing, BPM, code generation, design thinking, and SW education. Her research interests are in the areas of Software Visualization, Metaverse, AR/VR, Testing, BPM, Design thinking, SW education.

Ji Hoon Kong
https://orcid.org/0009-0004-1189-7674
e-mail : john.tooning@toonsquare.co

He worked at NCSoft Corporation from 2008 to 2014, followed by Samsung Electronics from 2014 to 2016, and Nexon Games from 2016 to 2019. He served as an Adjunct Professor in the Department of Game Media at Kaywon University of Art and Design from 2020 to 2021. He has been a Ph.D. candidate in the Software Engineering Lab at Hongik University since 2022. Additionally, since 2022, he has held the position of Director of Research at ToonSquare. His research interests include Game Software Engineering and 3D Interface Interoperability.

R. Young Chul Kim
https://orcid.org/0000-0002-2147-5713
e-mail : bob@hongik.ac.kr

He has been a Full Professor in the Department of Software and Communications Engineering at Hongik University since 2001. Since 2023, he has also served as the Director of the Metaverse SW Academy Institute. His research interests include Software Visualization, TMM (Test Maturity Model), Requirement Engineering, Metaverse, and AI-based Testing.