

International Conference on
Green and Human Information Technology 2024

ICGHIT 2024

Proceedings

Jan.23 - 25, 2024
Hanoi, Vietnam



S SsangYong
Information & Communications Corp.

NLT

<http://icghit.org/>



Applied Practices on Generating a Toon Image with Diverse Natural Language-based AI Tools

Young Sik Park
Software Engineering Laboratory
Hongik University
Sejong, Korea
park12160422@hongik.ac.kr

Woo Sung Jang
Software Engineering Laboratory
Hongik University
Sejong, Korea
uriel200@hongik.ac.kr

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

R. Young Chul Kim
Software Engineering Laboratory
Hongik University
Sejong, Korea
bob@hongik.ac.kr

Abstract—Currently we are researching UML model extraction based on natural language. But nobody researches a toon image creation through attribute identification from UML models. For comparison with better toon image generations, we are extracting AI images with diverse AI tools to give a guide on how to use AI image tools. We need to ensure whether the toon images created among these tools are creative, productive, and reusable or not. To do this, we propose the natural language-based image creation process for generating a toon image. Through this, we would like to suggest a way to easily create cartoons through natural language-based toon creation.

Keywords—Semantic roles extraction, Fillmore's case, Toon image generation, ChatGPT4, AI Tools

I. INTRODUCTION

When creating a toon image, we just need to analyze natural language-based sentences. To perform semantic analysis of sentences, we use "CHOMSKY & CHARLES J. FILLMORE's GRAMMER theory" to extract morpheme and semantic roles within a sentence in our approach [1,2,5].

Meanwhile, AI approach guys use LLM-based 'ChatGPT 4' and 'Bard' to extract semantic roles based on the same theory. For ChatGPT4, we find more than 30 semantic roles. For Bard, we find more than 13 semantic roles.

After the semantic role analysis is completed, a toon image is created with the semantic roles based on procedures within AI tools. The previous manual methods made it impossible to automatically create toon images.

II. RELATED WORKS

A. Fillmore's Case Grammar Theory

In Charles Fillmore's Case Grammar Theory, semantic roles (also known as "case roles") are conceptual categories that describe the relationship between a verb and the nouns or noun phrases in a sentence. These roles help in analyzing and understanding how different elements of a sentence interact with each other. Fillmore initially identified six core semantic roles:

1. Agentive (Agent): The entity that performs or causes the action. For example, in "The chef cooked the meal," "the chef" is the agent.
2. Instrumental (Instrument): The entity used by the agent to perform an action. In "She cut the paper with scissors," "scissors" is the instrument.

3. Dative (Recipient/Beneficiary): The entity that is the recipient or beneficiary of the action. In "I sent a letter to my friend," "my friend" is the dative
4. Factive (Fact): The entity that is either created or affected by the action. In "The storm damaged the house," "the house" is factive.
5. Locative (Location): The location where the action occurs. In "He sleeps in his bed," "his bed" is locative.
6. Objective (Goal/Object): The goal or object towards which the action is directed. In "She reads a book," "a book" is the objective.

Since 1971, the case grammar has been added to and modified by other theorists. TABLE 1 shows the definitions of the cases proposed by Fillmore & and other theorists [1,2,5].

TABLE I. ORIGINAL FILLMOER'S CASE

Case	Definition
Agent	A person or entity causing a verb's action to be performed.
Counter agent	The force or resistance against which a verb's action is carried
Object	An entity affected directly by a transitive verb's action.
Experiencer	A person or thing affected by a verb's action, replacing of the dative
Source	The place from which something moves.
Goal	The place to which something moves.
Locative	Location or spatial orientation of the state or action.
Instrument	The inanimate entity is causally involved in a verb's action.
Time	The time of the event

Fig 1 shows the refined Fillmore's Case Mechanism on how to adapt UML with the case mechanism for use-case extraction. This method analyzes semantic relationships between words based on the main verbs.

Therefore, processing natural language as requirement specifications may be more easily handled. Our refined case mechanism consists of nine cases: actor, counter-actor, object, theme, result, source, instrument, experience, and goal.

The process for generating toons from natural language can be seen in Fig 2. It is possible to create semantic roles with the existing methods and AI tools.

With ChatGPT4(DALL.E), we can passably make the image creation in next step.

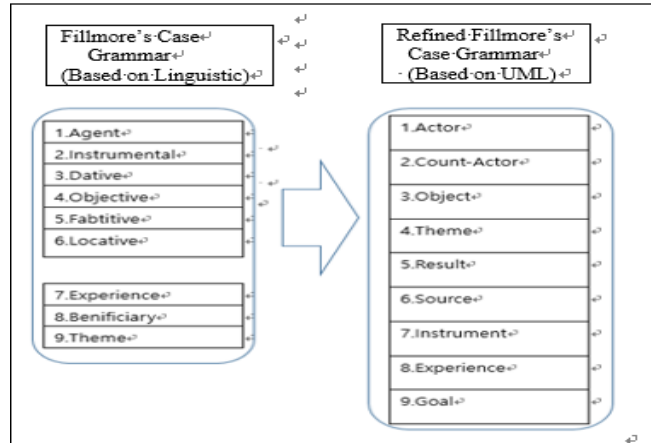


Fig. 1. Refined Fillmore's case mechanism [3,4].

III. OUR WORKS

A. A Toon Creation Process

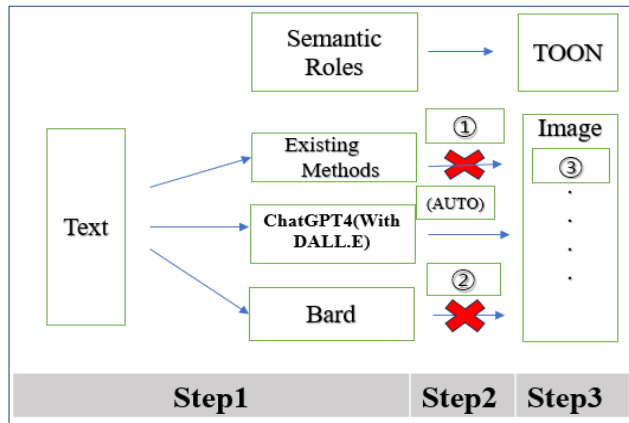


Fig. 2. Text to Toon (T_to_T) generation process with AI tools.

Step1. Create semantic roles with both the existing UML extraction methods and AI tools (such as ChatGPT4, and Bard)

Step2. Create an image using the created semantic roles as input. In ①of Fig 2, it is impossible to create images using the existing UML extraction method. As we expect, there are currently difficulties in image creation with Bard at ② of Fig 2. We find that only ChatGPT4(With DALL.E) can identify semantic roles and create fundamental images.

IV. AUTOMATIC MECHANISM WITH AI TOOLS

First step: Shows Natural Language-Based Sentences.

To generate toon images, give two sample sentences in Table 2.

In the first step in Fig 2, the extraction of semantic roles by ChatGPT4 (With DALL.E) is created for each stage, and the items that can be created at each stage are automatically extracted.

TABLE II. SAMPLE SENTENCES FOR GENERATING TOON IMAGES

Stage 1	"I saw a young girl with beautiful rose flowers meet the young and tall boy with a white face and blond hair who she loves in New York Park, located in New York City, at 11 o'clock today".
Stage 2	"After having a meal, they had tea time happily and drank coffee".

Second step : Extract Semantic Roles of This Sentence

TABLE III. SEMANTIC ROLES OF STAGES 1, AND 2 BY CHATGPT4 (WITH DALL.E)

Stage 1: "I saw a young girl with beautiful rose flowers to meet the young and tall boy with white face, blond hair who she loves in the New York Park that is located in New York City on 11 o'clock today"	
1. Experiencer: "I" - the person who saw the event.	
2. Agent: "A young girl" - the person actively carrying the flowers and intending to meet the boy.	
3. Theme: "Beautiful rose flowers" - what the girl carried.	
4. Purpose: "Meeting the "young and tall boy" - the reason for the girl carrying the flowers.	
5.Counter-Agent: "The young and tall boy" - the person the girl intends to meet.	
6.Location: "New York Park in New York City" - where the meeting is supposed to take place.	
7.Time: "11 o'clock today" - when the event took place.	
Stage 2: "After having a meal, they had tea time happily and drank coffee".	
1. Agent: "They" - likely referring to the young girl and boy, the ones performing the actions of having tea and drinking coffee.	
2. Theme: "Tea time" and "coffee" - the activities and items being consumed.	
3. Manner: "Happily" - describes the way the activity (tea time) was carried out.	
4. Time: Implied as "after having a meal" - indicating when the tea time and coffee drinking occurred (after their meal).	

Table 3 shows the PHASE related to the extracted semantic roles. The next shows key elements in the creation of visual storytelling media like comics, graphic novels, toons, and animation. Here's a brief explanation of each By ChatGPT4 With DALL.E):

1. **Setting:** This refers to the location and time where the story takes place. It includes the physical environment, cultural background, historical period, and any relevant societal norms or conditions. The setting provides context for the story and can significantly influence the plot and characters.
2. **Characters:** Characters are the individuals who populate the story. They can be humans, animals, mythical creatures, etc. Each character typically has their own distinct personality, background, motivations, and development arc. Characters drive the plot forward through their actions and interactions.
3. **Sequence:** In visual storytelling, sequence refers to the order and arrangement of scenes or panels. It's crucial for conveying the flow of events and the passage of time. A well-planned sequence ensures that the story is coherent and easy to follow.
4. **Dialogue:** Dialogue is the spoken words between characters. It's a key tool for revealing character traits, advancing the plot, and providing information. In comics or webtoons, dialogue is often presented in speech bubbles.
5. **Coloring and Shading:** Coloring involves applying colors to the artwork, which sets the mood and atmosphere of the scene. Shading is the technique of adding varying tones to create depth, volume, and a sense of light and shadow. Both are important for making the artwork visually appealing and dynamic.
6. **Detailing:** This refers to the addition of small, specific features in the artwork to enhance realism or convey more information about the setting or characters. Detailing can include background elements, textures, facial expressions, and intricate design elements.
7. **Final Touches:** The last step in the creation process, where the artist reviews the work and adds any necessary adjustments or enhancements. This can involve correcting errors, refining artwork, adding highlights, or making sure that the overall composition is balanced and complete

Third step: Generate A Toon Image

Fig 3 shows the toons implemented for each stage. We can not make a toon image in ChatGPT3.5 but can possibly make one in the recently announced ChatGPT4 (With DALL.E).

Additionally, due to complex considerations related to copyright, we may not reuse images at each stage.

V. CONCLUSIONS AND FUTURE RESEARCH

Based on natural language, both the existing UML model extraction method and the extraction results of semantic roles with AI Tools are almost similar. It is possible to proceed to make scenario creation and image production based on manual work with the semantic roles.

However, it is impossible to form an automatic creation mechanism for a toon image by manual work. Therefore, we attempt to create images through an automatic mechanism using recent AI Tools. The results of image creation with

ChatGPT4 (With DALL.E) are quite encouraging. We must have legal considerations such as copyright when images are created and used.

The collaboration of the existing UML model and the AI Tools is sufficient to give a fresh impact to us. In future research, we attempt to build mechanisms of toon production based on large quantities of sentences.

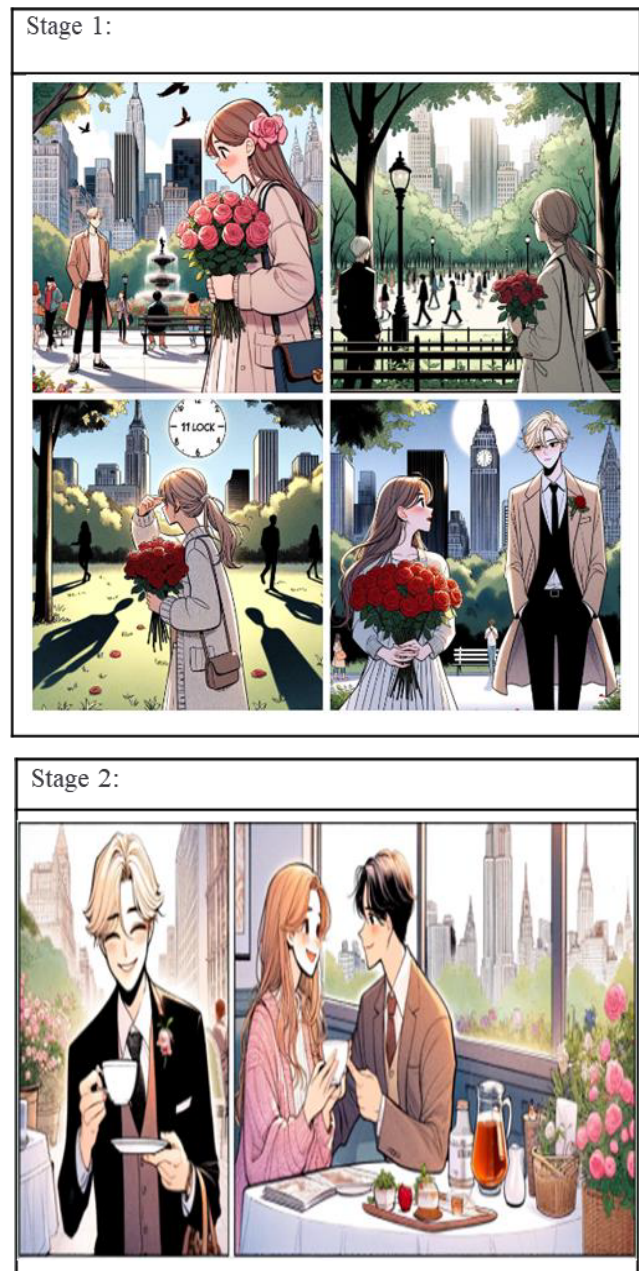


Fig. 3. Toon images of stages 1 and 2 by ChatGPT4(With DALL.E).

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: 100%).

REFERENCES

- [1] Fillmore C. J., "The Case for Case," Universals in Linguistic Theory, Bach&Harms Eds, New York: Holt, Rinehart&Winston, 1967.
- [2] Chikofsky, Elliot J., and James H. Cross. "Reverse engineering and design recovery: A taxonomy," Software, IEEE, 7(1), pp. 13-17, 1990.
- [3] Bo Kyung Park, R. Young Chul Kim "Effort Estimation Approach through Extracting Use Cases via Informal Requirement Specifications" Applied Sciences (SCIE), Vol 10 Issue 1-15, 2020 5. Anton, A.I. Goal-based Requirements Analysis. In Proceeding of the Second International Conference on Requirements Engineering (ICRE'96), Colorado Springs, CO, USA, 15-18 April 1996; pp. 136-144.
- [4] Park, B.K.; Son, H.S.; Kim, J.S.; Kim, R.Y.C. Refined Visual Modeling for Extracting Use Case Mechanism on Customer Requirements. In Proceedings of the International Conference on Convergence Technology ICCT2016, Jeju, Korea, 29, June 2016; Volume 6, pp. 640-641.
- [5] Fillmore, C.J. The Case for Case. Universals in Linguistic Theory, ed. by Emmon Bach and Robert T. Harms, 1-90. Holt, Rinehart & Winston: New York, 1968.

AI로 만들어가는 더 나은 세상

(주)비온시이노베이터가 한걸음 앞서 나아가겠습니다.

빅데이터

- 빅데이터 플랫폼 구축
- 데이터 자원 확보
- 데이터 품질 관리
- 정형/비정형 데이터
- 시각화

클라우드

- NHN 등 Cloud 파트너사
- G-Cloud Service
- Public Cloud Service
- Private Cloud Service
- Cloud Consulting

AI플랫폼

- 알고리즘을 통한 AI구현
- 머신러닝
- 데이터 라벨링
- AI서비스 개발

시스템통합

- 전문영역 SI (공공 부분)
- 시스템분석, 설계 및 개발
- 패키지소프트웨어 기반 시스템 개발
- 시스템관리, 운영 및 유지보수
- H/W, S/W 통합구축 및 관리

네트워크통합

- 네트워크 통합구축 및 유지관리
- 보안 환경 구축 및 유지관리
- 네트워크/보안 상태 점검 서비스
- 인프라 고도화 컨설팅

IT아웃소싱

- IT아웃소싱 컨설팅
- Network인프라운영서비스
- Application서비스
- 재해복구 · 보안관리서비스
- On-Site서비스

b BonC Innovators

(주)비온시이노베이터

대표전화 062)371-9339 팩스 062)371-9449

광주본부 광주광역시 북구 추암로 69(대촌동 959-51)

본 사 전라남도 나주시 그린로 229, 5층 501, 502호(빛가람동, 그린타워)

사람과 기술로 미래를 이끌어가는 기업
에스제이정보통신



고객의 믿음으로 성장하는기업 (주)에스제이정보통신이 더 나은 미래를 만들어갑니다.



SI	안정적인 시스템 구축 및 고객 업무환경에 맞는 최적의 솔루션 제공
ITO	다양한 경험과 기술로 고객의 기업가치의 경쟁력 향상
유지보수	고객의 업무혁신을 위한 다양한 분야의 종합 IT서비스 제공

SJ정보통신
SJinfo & communications

대전광역시 유성구 온구비로 8, 401호
(지족동 884-3, 보광빌딩)

TEL. 042-361-9990
FAX 042-361-9991

CONNECTING, SHARING & SHOWING

공간정보로 연결하고, 공유하고, 볼 수 있는 세상



 네오스펙트라

 올포랜드

 엘티메트릭

 명화지리정보

메타버스 플랫폼 서비스

솔루션 개발, SI

데이터 가치 창출

DB구축

공간정보서비스
Alliance

하늘 - 땅 - 바다 - 지하 - 메타버스 공간