

한국정보과학회
KOREAN INSTITUTE OF INFORMATION SCIENTISTS AND ENGINEERS

제26권 제1호
Vol. 26 No. 1



2024

제26회

한국 소프트웨어공학 학술대회 논문집

Proceedings of the 26th Korea Conference on
Software Engineering (KCSE 2024)

- 일시: 2024년 1월 31일(수) ~ 2월 2일(금)
- 장소: 강원도 평창 한화리조트(휘닉스파크점)

주최: 한국정보과학회, 한국정보처리학회

주관: 한국정보과학회 소프트웨어공학 소사이어티
한국정보처리학회 소프트웨어공학연구회

후원:  SOLUTIONLINK

(주)비트컴퓨터, (주)다한테크, 브이플러스랩(주),
슈어소프트테크(주), 한국정보통신기술협회,
(주)이에스지, T3Q(주)

자연어 기반 요구 사항을 활용한 UML 상태 다이어그램을 통해 만화 이미지 생성

진예진⁰¹, 공지훈², 김기두³, 장우성⁴, 김영철⁵

홍익대학교 소프트웨어공학연구실^{1,4,5}, 톤스퀘어², 한국정보통신기술협회³
 yejin_jin@g.hongik.ac.kr⁰¹, john.tooning@toonsquare.co², kdkim@tta.or.kr³,
 uriel200@hongik.ac.kr⁴, bob@hongik.ac.kr⁵

Generating a Cartoon Image via UML State Diagram with Natural Language-based Requirement Specs.

Ye Jin Jin⁰¹, Ji Hoon Kong², Ki Du Kim³, Woo Sung Jang⁴, R. Young Chul Kim⁵
 SELab., Hongik University^{1,4,5}, Toonsquare², Telecommunication Technology Association³

요약

현재 소프트웨어 산업에서는 자연어 기반의 요구사항 이슈가 대두되고 있다. 그러나 이러한 요구사항 분석의 정확도가 문제이다. 현재에도 요구사항 단계에서 오류가 가장 많이 발생한다. 자연어 기반의 요구사항 정의 및 분석이 절대적으로 필요한 시점이다. 이러한 문제를 해결하기 위해 자연어 기반 요구사항 분석에 언어학자 Chomsky와 Fillmore 이론의 언어학을 적용한다. 즉 형태소와 명사의 의미를 식별한다. 이러한 메커니즘을 바탕으로 자연어 기반의 만화 이미지를 생성을 제안한다. 이를 통해 만화 이미지 생성에 새로운 접근 방식을 제안한다.

1. 서론

본 연구는 인공지능 기반 사용자 대화형 멀티모달 인터랙티브 스토리텔링 3D 장면 제작 기술 개발에 관한 프로젝트의 성과를 담고 있다. 소프트웨어 개발 프로세스는 크게 요구사항 분석, 설계, 구현, 테스트, 유지보수로 나뉜다. 이 중 소프트웨어 오류의 66%가 시스템 명세 단계에서 나타난다[1].

본 연구에서는 요구사항을 명세하고 설계하는 과정에서 오류를 줄이기 위해 UML 모델을 사용한다. UML은 자연어에 비해 가시성이 높다는 장점이 있다. 자연어로부터 UML을 자동 추출하기 위해 촘스키의 구문 구조 분석을 통해 전체 문장의 구조를 파악하고, 형태소 단위로 분석한다. 분석된 문장 요소 간의 관계를 분석하기 위해 필모어의 의미역을 사용한다. 본 연구에서는 UML 중 상태 다이어그램을 추출하기 위해 언어학 기반의 추출 정보와 상태 다이어그램 생성에 필요한 항목을 매핑한다. 이를 통해 만화 이미지 생성이 가능한 Fabric.js를 활용하여 JavaScript 코드 템플릿을 구성한다.

본 논문의 구성은 다음과 같다. 2장은 연구와 관련된 어프로치에 대해 언급한다. 3장은 제안하는 메커니즘에 대해 언급하고 4장은 메커니즘을 적용한 사례 연구를 설명한다. 마지막으로 결론에 대해 언급한다.

2. 요구사항 모델링을 통한 상태 추출 어프로치

해당 연구는 시스템의 요구사항을 이해하기 위해 상태와 모드를 기반으로 자연어 요구사항을 모델링

한다[2]. 해당 연구에서 제안한 모델은 정해진 템플릿에 맞춰 자연어 요구사항이 표현된다. 이에 따라 모든 유형의 자연어 요구사항을 적용하기 어렵다는 단점이 있다.

본 연구에서는 언어학을 사용하여 복잡하게 구성된 요구사항을 체계적으로 분석한다. 규칙적인 템플릿의 제한 없이 사용할 수 있다.

3. 만화 이미지 생성 프로세스

본 연구에서 제안하는 자연어로부터 만화 이미지를 생성 프로세스는 그림 1과 같다.

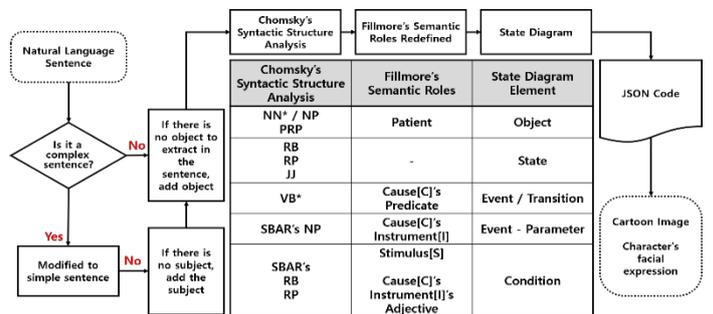


그림 1 자연어로부터 만화 이미지 생성 프로세스

자연어 문장이 입력으로 들어오면 복문을 단문으로 전처리한다. 버클리 파서를 통해 전처리 문장의 구문 구조를 분석한다. 분석된 구성 요소마다 재정의된 의미역을 부여한다. 언어학적 요소를 바탕으로 상태 다이어그램을 생성한다. 이 때, 상태 다이어그램의 상태는 구문 구조 분석 단계의 형태소와 매핑한다.

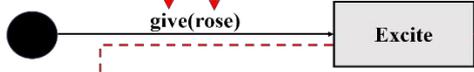
Requirement	<p>1) I saw Amy with beautiful rose flowers to meet the young and tall Tom with a white face and blond hair who she loves in the New York park that is located in New York City at 11 o'clock today.</p> <p>2) Amy gives roses to Tom, then Tom is excited</p> <p>3) While taking a walk in the park, they get hungry and go to a restaurant.</p> <p>4) Tom ordered food from the waitress.</p> <p>5) Tom happily received food from the waitress and fed it to Amy.</p> <p>6) After the meal, they happily drank tea and coffee.</p>	<p>1) I saw Amy who is carrying rose flowers. / Amy is going to meet Tom who is with a white face and blond hair. / Amy is going to meet Tom who she loves.</p> <p>2) Amy gives a rose to Tom. / Tom is excited.</p> <p>3) They walk in the park. / They get hungry. / They go to a restaurant.</p> <p>4) Tom ordered food from the waitress. / Tom received food from the waitress. / Tom was happy.</p> <p>5) Tom fed the food to Amy.</p> <p>6) After the meal, they happily drank tea and coffee.</p>
State Diagram		
Fabric.js & Cartoon Image	<pre data-bbox="279 526 774 692"><canvas id="c" width="200px" height="100px"></canvas> <script> var canvas = new fabric.Canvas('c'); var png="excite.png"; fabric.Image.fromURL(png, function(oImg) { canvas.add(oImg); }); </script></pre>	<p>코드 실행</p> 

그림 2 제안하는 프로세스 적용 사례 예시

매핑된 요소들을 바탕으로 상태 다이어그램을 추출한다. 추출된 상태는 만화 객체의 상태와 연결되고, 객체의 상태는 감정에 따른 표정의 변화로 이미지를 생성한다.

4. 적용 사례

본 연구에서 제안한 메커니즘을 만화 스토리에 적용하면 그림 2와 같다. 그림 2는 요구사항 문장을 분석하여 상태 다이어그램 일부와 이미지 생성을 위한 JavaScript 코드를 생성한다.

1) 자연어 분석

우선, 자연어 요구사항 문장을 단문 화 과정에 따라 재구성해야 한다. 기존의 두 번째 문장은 두개의 문장이 연결되어 있다. 따라서 이를 단문 화한다.

2) 버클리 파서를 통한 문장의 형태소 식별

단문 화된 문장은 촘스키의 이론 바탕으로 버클리 파서를 통해 형태소를 식별한다. 파서는 문장, 구, 절, 명사로 분석되고 POS 태그로 품사를 확인할 수 있다.

3) 형태소에 매핑되는 필모어의 역할 분석

품사 분석 결과를 통해 역할을 분석한다. 품사와 의미론적 역할을 매핑한 표를 바탕으로 역할을 유추하고, 주동사와의 관계와 의미에 따라 역할을 부여한다. 첫번째 문장의 동사 gives는 어떠한 감정의 변화를 일으킨 원인이기 때문에 Cause이고, 이 문장 안에 명사 rose는 이때 사용된 도구이기 때문에 Instrument이다. 두번째 문장의 주어 Tom은 상태의 변화를 겪기 때문에 Patient이고 이때, 형용사 excited가 변화된 상태가 된다.

4) 상태 다이어그램의 핵심 속성과 역할 매핑

기존 연구에서 만화 생성에 사용된 객체 다이어그램을 참고하여, 만화의 주요 객체를 식별한다[3]. 언어학 분석을 거친 문장의 각 요소는 상태 다이어그램의 요소와 매핑된다. 이를 바탕으로 주요 객체의 상태 다이어그램을 생성한다.

5) 상태 다이어그램 속성을 만화 속성과 매핑하여 만화 이미지 생성

주 객체의 감정이 나타나는 표정 에셋을 이미지 파일

로 호출하고 이를 캔버스에 나타나도록 한다.

5. 결론

본 연구는 자연어 요구사항 분석을 통한 상태 다이어그램 추출과 만화 이미지 생성 메커니즘을 제안한다. 제안된 메커니즘을 활용하면 체계적인 분석을 통해 만화 객체의 상태 변화를 자연스럽게 반영한 만화 이미지를 생성할 수 있다. 또한, 상태 다이어그램을 통해 만화 객체의 상태 변화를 명확하게 파악할 수 있다. 그러나, 주요 객체 하나의 상태 변화만을 나타내기 때문에 세부적인 객체의 상태는 확인할 수 없다.

향후 연구에서는 모든 객체의 상태를 다루고, 상세한 상태 다이어그램을 표현할 수 있도록 한다.

ACKNOWLEDGMENT

본 연구는 2023/2024년도 문화체육관광부의 재원으로 한국콘텐츠진흥원(과제명: 인공지능 기반 사용자 대화형 멀티모달인터랙티브스토리텔링 3D장면 저작 기술 개발, 과제번호: RS-2023-00227917,기여율:50%) 지원과 2023/2024년도 정부(교육부)의 재원으로 한국연구재단기초연구사업(과제명: NLP BERT Model 기반 자동 리팩토링을 통한 무결점 코드화 연구, 과제번호: No.2021R111A3050407,기여율:50%)의 지원을 받아 수행된 연구임

참고 문헌

- [1] Rupinder Kaur and Jyotsna Sengupta, "Software process models and analysis on failure of software development projects," IJSER, vol.2(2), 2012.
- [2] Yinling Liu, and Jean-Michel Bruel, "Modeling of Natural Language Requirements based on States and Modes," 2022 IEEE 30th International Requirements Engineering Conference Workshops, pp.190-194, 2022.
- [3] Jang Hwan Kim, and R. Young Chul Kim, "Cartoon Extraction Mechanism via UML Model based on Natural Language Requirement Specs.," 10th Annual Conf. on Computational Science & Computational Intelligence, 2023