

[제1회 크자회 시그니처대회]

Large Language Model을 활용한  
역사 교육 가상 현실 세계 플랫폼

2023년 9월 1일

Once Upon a Time

20213132 임 지 민 , 20215767 김 동 영

# 목 차

|                              |    |
|------------------------------|----|
| 연구 보고서 요약 .....              | 2  |
| <br>                         |    |
| I. 서론 .....                  | 3  |
| 1. 연구 동기와 목적 .....           | 3  |
| 2. 연구 방법 .....               | 3  |
| <br>                         |    |
| II. 본론 .....                 | 4  |
| 1. 이론적 배경 .....              | 4  |
| 가. Believable Behavior ..... | 4  |
| 나. 파이프라인 (Pipe Line) .....   | 4  |
| <br>                         |    |
| III. 결론 .....                | 7  |
| 구현 방법 및 연구 결과 .....          | 7  |
| 기대 효과 및 발전 방향 .....          | 9  |
| <br>                         |    |
| 참고문헌 .....                   | 10 |

<아이디어 보고서 요약>

# Large Language Model을 활용한 역사 교육 가상 현실 세계 플랫폼

Once Upon a Time

20213132 임 지 민 , 20215767 김 동 영

## 보고서 내용 요약

본 아이디어는 chat GPT와 가상 현실 기술을 결합하여 구현됩니다. 역사학적 자료 조사와 전문가들과의 협력을 통해 사용자가 선택한 시대의 사회, 문화, 건축물 등을 구체적으로 설계합니다. 사용자는 연도와 특정 역사적 상황, 장소 등을 입력합니다. 먼저 Large Language Model을 통해 맵 정보와 맵 자체를 생성합니다. 이후 시스템에서 사용자의 입력과 생성된 맵 정보, 연도와 역사적 상황들을 고려하여 generative agent의 seed memory를 생성하면 environment 생성이 끝납니다. 사용자는 이렇게 생성된 사회 안에서 NPC 과 상호작용하며 영향을 끼칠 수 있습니다. 사용자와 세계와의 상호작용을 통해 얼마든지 세계는 변화할 수 있으며, 새로운 인공사회가 형성될 수 있습니다. 사용자의 주체적인 변화가 영향을 미칩니다. 구체적으로 설정해야 세계의 완성도가 높아집니다. 설정하지 않은 부분에 대해서는 오류가 존재할 수 있습니다.

# I. 서론

## 1. 연구 동기와 목적

이 프로젝트의 동기와 목적은 사용자에게 역사와 문화에 더 가까이 다가갈 수 있는 새로운 경험을 제공하고자 함에 있습니다. 대화 기반 가상 현실 시스템을 통해 특정 시대의 역사와 문화를 체험하며 사용자들의 역사적 이해를 촉진하고 창의적인 역사 체험을 가능케 하는 것을 목적으로 합니다.

## 2. 연구 방법

논문 Generative Agents: Interactive Simulacra of Human Behavior를 참고하여 연구를 시작하였습니다. 연구의 바탕이 되는 이 논문을 소개하자면, 생성 에이전트는 대화형 애플리케이션을 위한 인간 행동의 믿을 수 있는 시뮬레이션입니다. 이 작업에서 우리는 The Sims를 연상시키는 샌드박스 환경을 25개의 에이전트로 채워 생성 에이전트를 시연합니다. 사용자는 상담원이 하루를 계획하고, 뉴스를 공유하고, 관계를 형성하고, 그룹활동을 조정하는 과정을 관찰하고 개입할 수 있습니다.



[ 그림 1 ] 생성 에이전트

## II. 본론

### 1. 이론적 배경

\* 논문 제목: Generative Agents: Interactive Simulacra of Human Behavior

#### 가. Believable Behavior

Large Language Model을 사용하여 신뢰할 수 있는 인간 행동을 모방하도록 했습니다.

##### - Agent Avatar and Communication

논문에서는 25명의 unique agent를 생성합니다. 각 agent의 자아는 그들의 occupation과 타 agent들과의 relation을 설명하는 한 문단의 language description으로 기술되며 이는 seed memory(memory stream의 첫 번째)에 위치합니다.

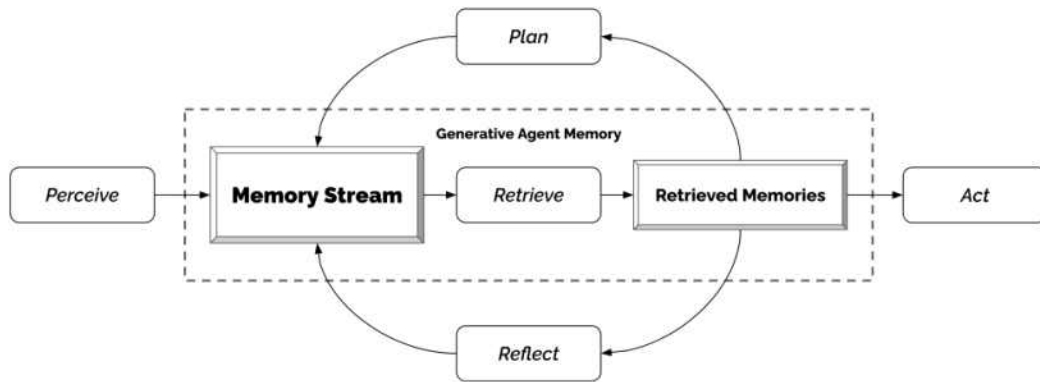
또한 user control이 가능합니다. end user는 자연어를 통해 agent들과 상호작용할 수 있습니다. 그리고 inner voice라는 방식으로 agent를 조종할 수 있습니다.

##### - Environmental Interaction

end user와 agent는 이 세계의 object의 상태에 영향을 줄 수 있습니다.

#### 나. 파이프라인 (Pipe Line)

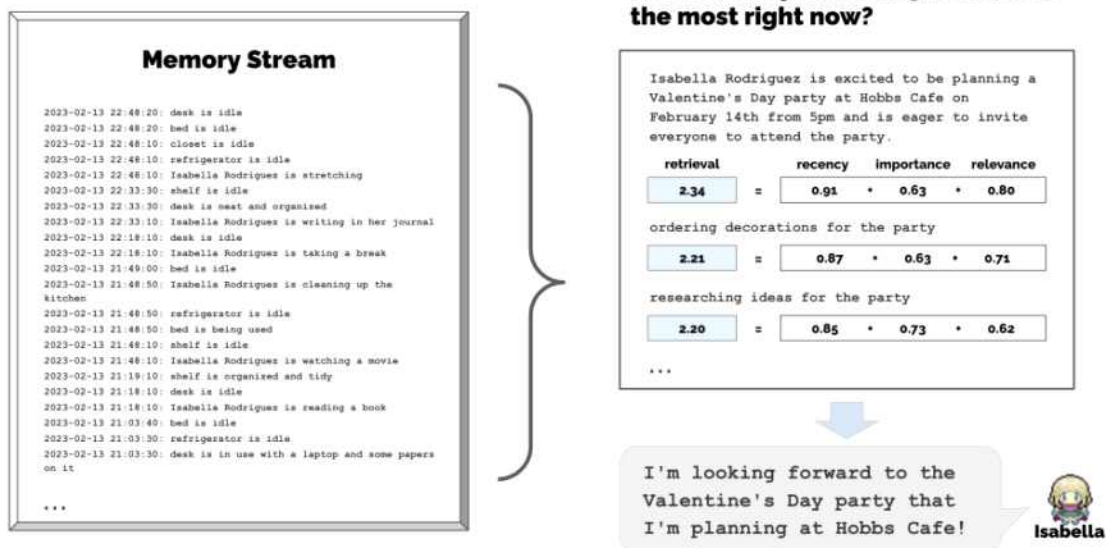
이를 가능하게 하려면 memory stream과 retrieval, reflection, planning 4단계의 과정을 거치는 파이프라인을 생성합니다. 생성 에이전트 아키텍처 (파이프라인) 구조는 다음 [그림 1]과 같습니다.



[ 그림 2 ] 생성 에이전트 아키텍처 (파이프라인) 구조

## - Memory and Retrieval

Memory stream은 기본적으로 관찰(observation)에 기반하여 저장됩니다. retrieval function은 agent의 현재 상황을 입력으로 받고 memory stream의 subset을 출력으로 내놓는 language model입니다. recency와 importance, relevance를 고려합니다. recency(exponential decay function)는 얼마나 최신 정보인가를 의미합니다. importance는 얼마나 중요한 정보인가를 의미합니다. large language model을 사용해 점수를 산출하며, 논문에서 프롬프트가 있습니다. relevance는 현재 상황과 얼마나 관련 있나를 의미합니다. 각 memory 간의 embedding vector와 query vector의 embedding vector 간의 cosine similarity 값을 통해 산출됩니다.



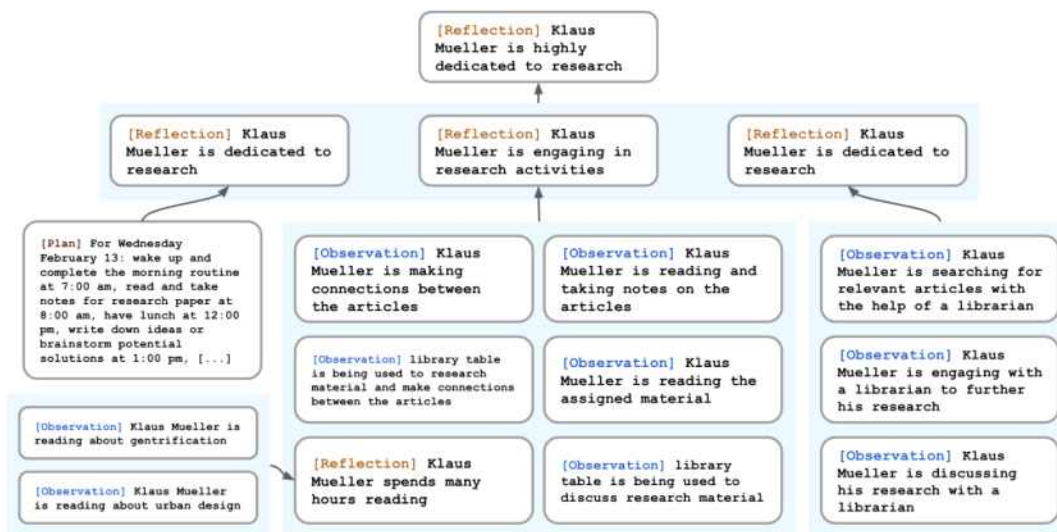
[ 그림 3 ] 메모리 스트림

## - Reflection

Reflection은 Observation보다 상위계층의 memory입니다. 상위계층일수록

생성 에이전트에 의해 생성되는 생각들의 추상화 정도가 커집니다. importance score의 합계가 150점이 넘을 때마다 reflection 작업을 실행합니다. 하루에 2, 3 번 정도가 됩니다.

먼저 language model에 이 질문들을 query로 사용해서 retrieval 작업을 실행해서 relevant memories 를 뽑아냅니다. 최근 100개의 record를 넘긴 뒤 “Given only the information above, what are 3 most salient high-level questions we can answer about the subjects in the statements?”라는 질문을 통해 candidate question을 산출합니다. 그 relevant memories 들을 사용해서 high-level insight를 Large Language Model을 통해 뽑아냅니다. 이를 memory stream에 추가합니다. Reflection은 observation뿐만 아니라 다른 reflection들을 사용해서 이뤄집니다.



[ 그림 4 ] Klaus Mueller의 반사 트리

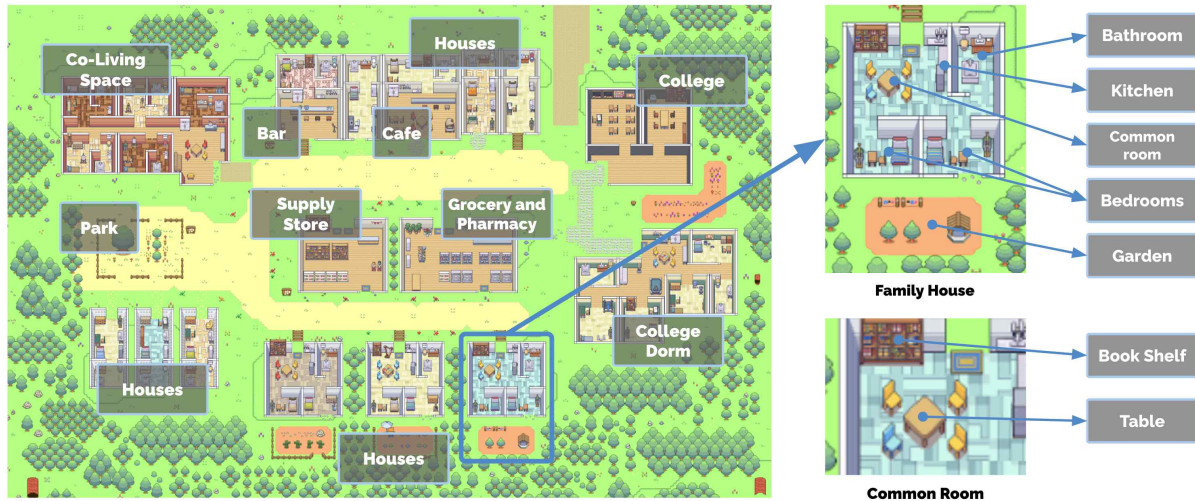
#### - Sandbox Environment Implementation

Phaser web game development framework를 사용했습니다. agent avatar와 environment map, collision map을 통해 구현됩니다. generative agent 들이 sandbox environment에 영향을 줄 수 있도록 추가로 구현했는데, 서버가 각 agent의 현재 위치, 행동, sandbox object 들 간의 상호작용을 저장하는 JSON 파일을 유지하도록 했습니다. 각 time-step마다 서버가 JSON 파일을 파싱해서 각 에이전트를 새로운 장소로 옮기고 object 들의 상태를 변경합니다.

### III. 결론

#### 1. 구현 방법 및 연구 결과

본 아이디어는 chat GPT와 가상 현실 기술을 결합하여 구현됩니다. 역사학적 자료 조사와 전문가들과의 협력을 통해 사용자가 선택한 시대의 사회, 문화, 건축물 등을 구체적으로 설계합니다.



[ 그림 5 ] 영역에 라벨이 표시된 스몰빌 샌드박스 세계

사용자는 연도와 특정 역사적 상황, 장소 등을 입력합니다. 먼저 Large Language Model을 통해 맵 정보와 맵 자체를 생성합니다. 이후 시스템에서 사용자의 입력과 생성된 맵 정보, 연도와 역사적 상황들을 고려하여 generative agent의 seed memory를 생성하면 environment 생성이 끝납니다. 사용자는 이렇게 생성된 사회 안에서 NPC 과 상호작용하며 영향을 끼칠 수 있습니다.



그림 6



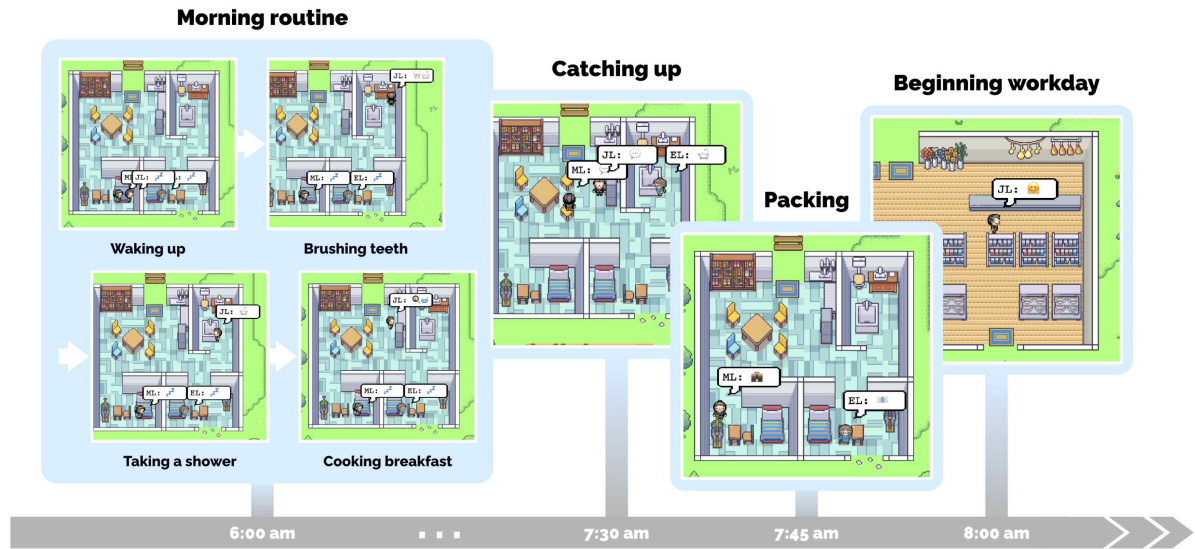


그림 7

| 서식  |
|---|
| <ul style="list-style-type: none"> <li>- 년도:</li> <li>- 장소:</li> <li>- 사건: 선택하지 않으면 당시의 큼직한 사건 중 무작위로 생성합니다</li> <li>- 사람의 페르소나, NPC 수, NPC 직업, 등등</li> <li>- 텍스트로 체험 <ul style="list-style-type: none"> <li>- 책의 내용을 추가할 수도 있습니다.</li> </ul> </li> </ul> |

사용자와 세계와의 상호작용을 통해 얼마든지 세계는 변화할 수 있으며, 새로운 인공사회가 형성될 수 있습니다. 사용자의 주체적인 변화가 영향을 미칩니다. 구체적으로 설정해야 세계의 완성도가 높아집니다. 설정하지 않은 부분에 대해서는 오류가 존재할 수 있습니다.

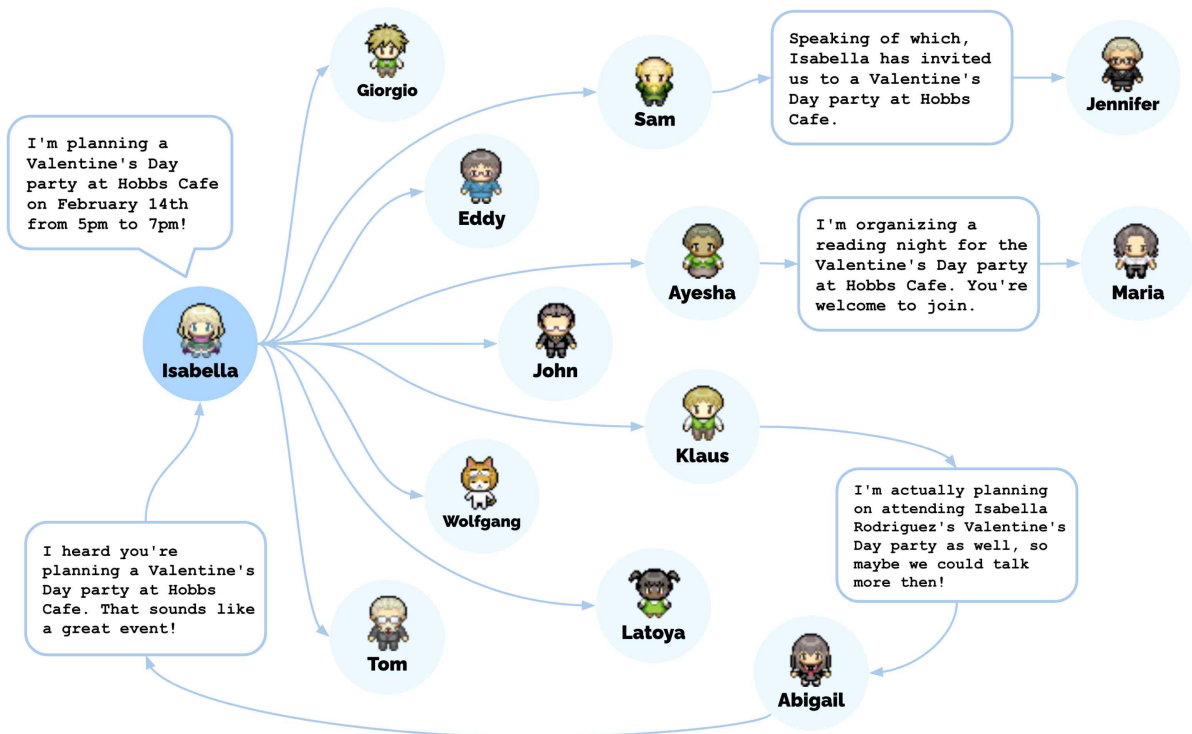


그림 8

## 2. 기대 효과 및 발전 방향

### - 역사 교육에서의 혁신

교육 분야에서 혁신적인 역할을 할 것으로 기대됨. 학생들은 역사와 문화를 가상 현실 환경에서 체험하면서 학습하고, 이를 통해 더 깊은 이해와 학습 동기부여를 얻을 것임.

### - 게임화 (Gamification)

사용자가 역사에 개입하여 역사적 사건들의 흐름을 바꿀 수 있다는 점을 통해 역사 교육을 게임처럼 만들어 학생들의 흥미를 끌 수 있으며 역사 공부에 대한 거부감과 진입장벽을 낮출 수 있습니다.

### - 사용자들 간의 경험 공유 플랫폼

본 프로젝트는 사용자가 특정 연도와 시대적 사건, 장소, description 등을 입력하면 그것에 맞게 Large Language Model이 맵 정보와 agent 정보를 생성하는 방식으로 구현됩니다. 사용자들은 본인들의 경험과 프롬프트들을 공유할 수 있습니다. 이는 사용자들 간의 역사적 식견 공유의 장으로 작용하고 새로운 역사 연구의 장을 열 수도 있습니다.

- Baseline 성능 향상에 따른 사용자 효용 증가

본 프로젝트의 사용자 만족감은 Chat GPT 같은 large language model의 성능에 따르기 때문에 baseline의 성능이 올라갈수록 사용자 효용이 증가할 수 있음. large language model은 학계와 산업계에서 많은 관심을 받고 있고 그에 따른 연구도 활발히 진행 중이므로 관련된 연구가 진행될수록 완성도가 높아질 것으로 추정됩니다.

## 참고문헌

Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative Agents: Interactive Simulacra of Human Behavior. In The 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23), October 29-November 1, 2023, San Francisco, CA, USA. ACM, New York, NY, USA, 22 pages. <https://doi.org/10.1145/3586183.3606763>