Table 6: Comparison of different LLM role-playing approaches.

Methods	Fine-tuning Based	Retrieval Based	Plugin Model
LLM Training	YES	No	YES
Character Data Labeling	YES	No	YES
Computational Burden	High	Low	Moderate
Character Data Organization	All characters shared	One character, one corpus	One character, one plugin
Adaptation to Unseen Roles	Hard	Easy	Hard
Modifying LLMs' Knowledge	Hard	Easy	Hard

D   Additional Experiments

D.1   Word Clouds

In this experiment, we collect all responses generated by GPT-4o-mini when role-playing characters from the Harry Potter series and visualize them using word clouds to highlight frequently used terms. Larger words in the figures indicate higher frequency. As shown in Figure 5, when playing as Harry Potter, the model emphasizes terms such as friends, Hogwarts, Hermione, Voldemort, and brave. In contrast, Figure 6 shows that when role-playing Voldemort, the most frequent terms include world, power, death, and dark.

D.2   Demonstration of RoleRAG retrieval

Figure 7 illustrates the types of information retrieved by RoleRAG in response to an interview question posed to an LLM role-playing Ludwig van Beethoven. The question asks about Beethoven's relationship with Haydn and Mozart. RoleRAG first identifies the entities in the query, assesses their relevance to Beethoven, and determines their specificity.

Since all three are specific entities related to Beethoven, we locate them in the knowledge graph using semantic matching and the name mapping table. We then retrieve detailed information about each entity and their relationships with Beethoven. Finally, the retrieved context is aggregated and provided to the LLM to answer the user's question.

E   Prompts in our experiments

This section contains the prompts used in our experiments. Figures 8, 9, 10, and 11 show the prompts used for generation and scoring. Green text in curly braces represent text that is replaced based on the context.

Demonstration of RoleRAG retrieval step

Question: {Beethoven, What was the nature of your relationship with Haydn and Mozart?}

Entity Extraction: {entity type, name, familiarity with Beethoven (reason, YES|No), Level (specific|general)}
(character<|>Beethoven<|>Beethoven must know himself<|>Yes<|>specific)
(character<|>Joseph Haydn<|>Haydn was Beethoven's teacher and a significant influence on his early compositions<|>Yes<|>specific)
(character<|>Wolfgang Amadeus Mozart<|>Mozart was a contemporary of Beethoven and a major influence on his work, although they did not have a direct relationship<|>Yes<|>specific)

Retrieved Content:
—- Entity Information——
BEETHOVEN,CHARACTER, "Ludwig van Beethoven was a renowned German composer and pianist who significantly contributed to the landscape of classical music, transitioning between the late Classical and early Romantic periods. Born in Bonn, he moved to Vienna, ... "

JOSEPH HAYDN,CHARACTER,"Joseph Haydn was a prominent composer who played a significant role in Beethoven's musical education, particularly in the area of counterpoint. He was an esteemed figure in the Classical music genre and served as a pivotal influence during Beethoven's early years in Vienna, introducing him to various musical concepts and techniques."

MOZART,CHARACTER,"Wolfgang Amadeus Mozart was a prominent and renowned composer whose significant success and legacy played a crucial role in influencing Johann, particularly in promoting Beethoven as a child prodigy. His impact extended deeply into Beethoven's formative years as a composer, shaping the young musician's development and artistic direction."

—- Relation Information——
BEETHOVEN,JOSEPH HAYDN, "Beethoven, an influential composer in the history of classical music, studied under the guidance of Joseph Haydn, a prominent figure known for his significant contributions to the development of the symphony and string quartet. Under Haydn's direction, Beethoven acquired essential skills in counterpoint and composition, which greatly influenced his early works. Additionally, Haydn played a crucial role in introducing Beethoven to a broader musical community and providing mentorship in Vienna, marking a pivotal development in Beethoven's career. This relationship between the two composers highlights the impact of mentorship and collaboration in the evolution of classical music."

BEETHOVEN,MOZART,"Beethoven, a prominent composer, was notably influenced by the style of Mozart, which played a significant role in his artistic development. This relationship highlights the profound impact that Mozart's musical elements had on Beethoven's compositions."

Figure 7: Use case of retrieval step in our RoleRAG.

Prompt for Generating Knowledge Exposure Scores

Play act as {character}, who is {description}. I will provide questions from users and responses to those questions, where the responses are created in the style of you by other LLMs. You are required to judge and assess whether the response to a user's question match the knowledge and experience of you. To evaluate the response, consider the following aspects:

(1) Whether stories/events discussed occurs in the same period as you.
(2) Whether objects in the response have relevance to you.
(3) Whether locations in the response are correct in your experience.
(4) Whether persons mentioned have accurate relationships with you.

Based on the given instructions, provide a brief analysis of the given response. Then rate the response using a single score from 1 to 10, where a higher score indicates greater consistency with your knowledge.

Please provide your output in the following format:
Analysis: <analysis>
Rating: <rating>

Test Begin
User Question: {question}
Response: {response}
Output:

Figure 8: The prompt used for generating knowledge exposure scores.

Prompt for Generating Knowledge Hallucination Scores

Play act as {character}, who is {description}. I will provide questions from users and responses to those questions, where the responses are created in the style of you by LLMs. Based on your knowledge and experience, you must judge and assess whether the response to the question contains hallucination (fabricated or incorrect information). To evaluate hallucination, consider the following aspects:

(1) Whether the events, objects, locations, or persons mentioned are consistent with your established story and background. A response that is not consistent with your lore is considered as hallucination.

(2) Whether the response demonstrates a deep level of knowledge about a topic or concept that does not make sense for you to have, due to factors such as the topic not existing in your time period or universe. A response may refer to a topic if the question directly asks about the topic. However, answering the question with great detail would be considered as hallucination.

Based on the given instructions, provide a brief analysis of the given response. Then rate the response using a single score from 1 to 10, where a score of 1 represents a response with no hallucination and a 10 represents a response that does not make sense at all.

Provide your output in the following format:
Analysis: <analysis>
Rating: <rating>

Test Begin
User Question: {question}
Response: {response}
Output:

Figure 9: The prompt used for generating knowledge hallucination scores.