Towards Enhanced Immersion and Agency for LLM-based Interactive Drama

Anonymous ACL submission

Abstract

LLM-based Interactive Drama is a novel AI-based dialogue scenario, where the user (i.e. the player) plays the role of a character in the story, has conversations with characters played by LLM agents, and experiences an unfolding story. This paper begins with understanding interactive drama from two aspects: **Immersion**—the player's feeling of being present in the story—and **Agency**—the player's ability to influence the story world. Both are crucial to creating an enjoyable interactive experience, while they have been underexplored in previous work. To enhance these two aspects, we first propose *Playwriting-guided Generation*, a novel method that helps LLMs craft dramatic stories with substantially improved structures and narrative quality. Additionally, we introduce *Plot-based Reflection* for LLM agents to refine their reactions to align with the player's intentions. Our evaluation relies on human judgment to assess the gains of our methods in terms of immersion and agency.

1 Introduction

Human existence is deeply rooted in emotional expression and connections. AI-based dialogue systems can provide instant and cheap companionship, where humans are free to express themselves. However, such open-ended interactions fall short of building a deep and complete emotional connection, which naturally emerges from shared experiences. This gap can be bridged by co-constructing an enjoyable interactive experience between humans and AI. This paper studies *LLM-based Interactive Drama*, a new format of dialogue application, which allows the human player to experience a story with dramatic development.

Drama is a classical mode of storytelling and emotional expression, which utilizes dialogues to tell the story. Compared to classical drama, where the audience passively watches characters having conversations, interactive drama (Mateas, 2000;
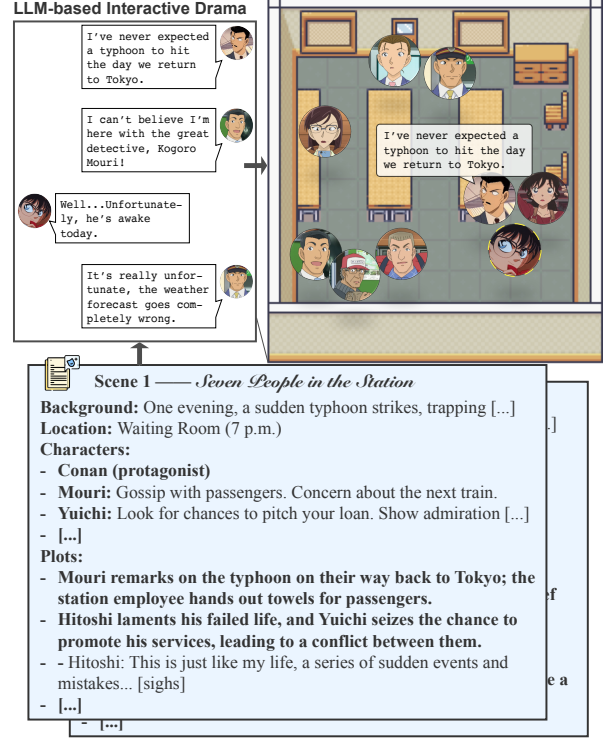

Figure 1: A demonstration of LLM-based interactive drama. Rather than the imaginary scene on the left, our work studies the dialogue-form drama on the right. Below is the drama script, adapted from *Detective Conan*.

Mehta et al., 2007) transfers the audience into a "player", who is able to interact directly with the characters and experience the story from a first-person point of view. In LLM-based Interactive Drama, the story is performed by LLM agents (Brown et al., 2020; OpenAI, 2023; Dubey et al., 2024; Jiang et al., 2023). Figure 1 offers a demonstration. Guided by a drama script, the LLM agents simulate characters (e.g. *Mouri*, *Yuichi*) and make a conversation, while the player (playing as *Conan*) interacts with them. Unlike traditional drama scripts, where every line is scripted, the drama script for LLM agents is only defined by a sequence of plots. Following the plots, they generate
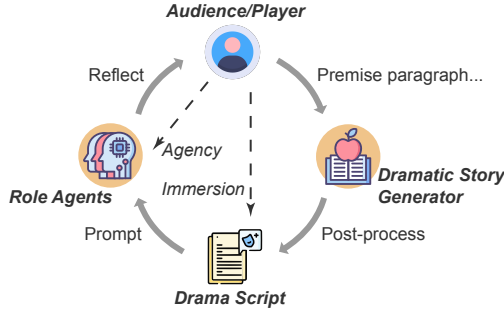
Figure 2: The proposed immersion-agency paradigm for LLM-based interactive drama.

dialogues and instantaneous interactions with the player. However, traditional interactive experiences are often rigid and constrained by predefined plots. Therefore, we seek to harness LLMs' capabilities in dialogue and character simulation to break that limitation, building deeper emotional connections with the player. In this paper, we study interactive drama in the dialogue format and do not discuss scenery generation techniques.

To understand the interactive experience enabled by LLM-based interactive drama, we introduce the two key aspects discussed in Mateas (2000): **Immersion** and **Agency**. While previous work focuses on generic architectures (Wu et al., 2024; Han et al., 2024), it hasn't carefully discussed these two aspects, as well as the relationship between them.

Immersion refers to the sensation of being present in the story. In addition to visual assets and music, this is fundamentally determined by the quality of the story as well as the storytelling process. Any missing plot or interruption can break the player out of the story. Moreover, *immersion can be further enhanced when the player is able to experience the story of their own making.* This is because the self-made character may have a closer connection to the player themselves, making it easier for them to immerse in the story.

Agency refers to the ability to influence the story world. Agency can be straightforwardly understood as engagement. Traditionally, the player is provided with some options and specific choices can lead them to different plot branches. However, such agency has been determined by the script. In LLM-based interactive drama, the player is endowed with more freedom by being able to have open-ended interactions with characters. We thus focus on a more general aspect of agency—the characters can exhibit meaningful transformation of reaction induced by the player's behavior. For

example, a player might attempt to extract a secret from a non-player character (NPC), seek temporary companionship, or advance the plot in a specific way. The NPCs should understand the player's intentions and offer meaningful feedback.

The above theory provides foundation for our work. We introduce the immersion-agency paradigm for LLM-based interactive drama in Figure 2. To enhance immersion and agency based on this paradigm, we make two key contributions.

• The ideal vision is that players are able to create a complete drama script effortlessly and then immerse in it. This duty can be delegated to an LLM generator, which crafts the dramatic story based on a player-specified premise paragraph. However, we find the stories generated by current methods lack basic dramatic structures and compelling conflicts, leaving a significant gap to human-written ones. To mitigate this gap, we propose *Playwriting-guided Generation* in Sec. 3.

• Previous work has largely overlooked agency in constructing LLM-based role agents. In Sec. 4, we present *Plot-based Reflection*, which adapts and creates plots by analyzing player-related memories to better guide characters' behavior. A hybrid architecture is then introduced to optimize computational efficiency.

2 Related Work

Interactive drama driven by computer systems has been proposed for quite a long time (Bates et al., 1991; Laurel, 1993). More recently, Wu et al. (2024) first discuss LLM-based interactive drama, defining its six primary elements. The director-actor architecture is first used in Han et al. (2024), while Magee et al. (2024) model a single character with an ego and a superego to capture its development. In our work, we employ a director agent to program the plots periodically. Wu et al. (2024) propose a curriculum learning method to fine-tune the role agents. Our work, in parallel, proposes a comprehensive framework suitable for any LLMs.

AI-Generated Stories It has been widely used in recent years to scale the training performance by using AI-generated data (Wang et al., 2023b; Lee et al., 2023; Wu et al., 2024). Our research focuses on the use of LLMs to assist humans in creating high-quality literature, such as poetry (Ormazabal et al., 2022; Chakrabarty et al., 2022), music (Hernandez-Olivan et al., 2022; Zhu et al., 2023), stories (Tan et al., 2021; Tian et al., 2024). Our

2

Love	Phoenix	Cinderella	Love Triangle
Love	*Phoenix*	*Cinderella*	*Love Triangle*
- The protagonist meets their heart's desire.	- The protagonist begins at a low point in life or society.	- The protagonist is initially in a state of hardship.	- The protagonist is torn between two admirers.
- A romantic relationship blooms but is soon met with obstacles.	- Seize a key opportunity and face setbacks along the way.	- Receive an opportunity to escape, hindered by societal barriers like lineage.	- The dynamic creates competition and jealousy among the three.
- The lovers overcome the obstacles or get separated. *"Titanic"*	- Achievement of life goals above the struggles. *"The Great Gatsby"*	- Virtue and talents are recognized towards a happy life. *"Jane Eyre"*	- A choice is eventually made solving the situation. *"The Twilight Saga"*
Revenge	*Family*	*Reunion*	*Savior*
- The protagonist suffers great harm or betrayal.	- Complex relationships between family members.	- The protagonist leaves familiar places for a reason.	- The protagonist, faced with a responsibility or a call, decides to forge ahead.
- They devise and execute a meticulous plan.	- Tensions arise either from the family or external forces.	- They grow through a series of trials and tribulations.	- Break through difficulties and even life threats.
- Attainment of satisfaction or a lingering sense of void. *"Hamlet"*	- Understanding and deepened affections. *"Little Women"*	- They return home and reunite with loved ones. *"Odyssey"*	- Eventually save the people. *"Dune"*

Table 1: Eight dramatic situations. We describe them using: - Setup, - Confrontation, - Resolution.

method generates a story given a short paragraph. This is more challenging than generating the drama script given a complete story (Zhao et al., 2024).

Simulating Dramatic Characters with LLM Agents Recent years emerges many companion-oriented AI applications such as `character.ai`, where LLM agents are developed to simulate various characters (Shanahan et al., 2023; Shao et al., 2023; Chen et al., 2024; Lu et al., 2024). A new request arises, players are not satisfied with aimless conversations, but would like to go through a story, an adventure with AI, indicating the very need for our research, which connects NLP with research in psychology (Magee et al., 2024), art (Mirowski et al., 2023), and narratology (Todorov, 1969). Dramatic characters evolve over time, undergoing shifting inner motivations, which greatly distinguishes them from conversational agents.

3 Playwriting-Guided Generation

The dramatic story generator seeks to craft an engaging story with dramatic development based on a player-specified paragraph, which we call the premise paragraph. It includes the background, protagonist description, and beginning of the story.

However, recent studies have shown that LLMs' storytelling capabilities fall significantly short of human standards (Tian et al., 2024; Chakrabarty et al., 2024). This phenomenon is contradictory to their pre-training process, during which LLMs are exposed to a vast amount of human assets, including numerous masterpieces of literature.

Indeed, **being a brilliant playwright demands not only exposure to great works but also mastery of sophisticated playwriting techniques.** We conjecture that most LLMs' fine-tuning processes lack emphasis on the latter dimension. As a result, while LLMs accumulate a large repository of writing samples, they struggle to distill and apply the core principles from them effectively.

In this section, we present a series of playwriting techniques and propose *Playwriting-Guided Generation*, a method that integrates playwriting techniques as heuristic principles to guide the LLM generator in crafting more engaging and structured dramatic stories. Our method turns out to significantly improve the LLMs' storytelling quality. However, it doesn't endow LLMs with human-level power. Humans' emotions and real-life experiences enable us to create works of greater depth and resonance.

3.1 Dramatic Situations

Polti et al. (1945) identified thirty-six dramatic situations by analyzing a vast collection of literary works. Dramatic situations can be straightforwardly understood as the structure a dramatic story unfolds. More recent studies refine and condense these situations into a more streamlined set (Baker, 2022). These dramatic situations are more adapted to modern stories, covering vast majority of dramatic conflicts and human emotions. We show them in Table 1. To provide a clear understanding, we illustrate each using Aristotle's three-act structure (Field, 2005), which divides a drama into setup, confrontation, and resolution.

3.2 Narrative Techniques

Dramatic situations can be regarded as macro-level playwriting techniques, offering a framework for

	Suspense	Twist	Non-Linear
Qwen2.5-72b	4 (8%)	3 (6%)	1 (2%)
GPT-4o	6 (12%)	3 (6%)	0
Our Method	21 (21%)	37 (74%)	14 (28%)

Table 2: Narrative techniques used in 50 LLM-generated stories. Our method is based on GPT-4o.

Algorithm 1 Playwriting-Guided Generation

Input: writer LLM $\mathcal{A}$, critic LLM $\mathcal{B}$, premise paragraph $\mathbf{w}$, dramatic situations $\mathcal{D}$, narrative techniques $\mathcal{N}$
Output: dramatic story $\mathcal{S} = \{s_k^*\}$
1: **for** $i = 1$ to 3 **do**
2:      Sample 1 dramatic situation d_i from $\mathcal{D}$
3:      Sample 3 narrative techniques $n_i^{1\sim3}$ from $\mathcal{N}$
4:      Generate the story $\mathcal{S}_i = \mathcal{A}(\mathbf{w}, d_i, n_i^{1\sim3})$
5:      Comment the story $c_i = \mathcal{B}(\omega_i, d_i, n_i^{1\sim3})$
6:      Revise the story by the comment $\mathcal{S}_i \leftarrow \mathcal{A}(\mathcal{S}_i, c_i)$
7: **end for**
8: Vote for the best story $\mathcal{S} = \{s_k\}$ from $\mathcal{S}_{1\sim3}$
9: **for** $i = 1$ to 3 **do**
10:      Refine the story with finer details $\{s_k^*\} \leftarrow \mathcal{A}(\{s_k\})$
11: **end for**

---

constructing a logical and efficient story structure. In this part, we further focus on micro-level narrative techniques. We summarize six widely-used narrative techniques, drawing from Archer (1913) as well as a wide range of modern storytelling mediums (e.g. web novels, games). They are **Suspense**, **Twist**, **Non-linear Narrative**, **Multiple Narrative**, **Irony**, and **Symbolism**. Details of each technique can be found in Appendix A.

We examine the use of three representative narrative techniques in two state-of-the-art LLMs (GPT-4o (OpenAI, 2023) and Qwen2.5-72b (Yang et al., 2024)). Table 2 compares the vanilla prompting method to our method (will be detailed below) given 50 premise paragraphs. It suggests that, without explicit prompting, both Qwen2.5 and GPT-4o are almost not able to apply any type of narrative techniques in their storytelling.

3.3 Story Generation

The entire story generation process is shown in Algorithm 1. The generator we use is GPT-4o and we put the prompts in Appendix A. First, we initialize the sets of dramatic situations $\mathcal{D}$ and narrative techniques $\mathcal{N}$. From them, we uniformly sample one dramatic situation and three narrative techniques (line 2~3). We represent a story $\mathcal{S}$ as a sequence of narrative sentences $\{s_1, s_2 \cdots\}$, briefly denoted as $\{s_k\}$. We prompt the writer LLM $\mathcal{A}$ to craft the story, applying the selected dramatic situation and all three narrative techniques.

Sampling The first challenge is to determine which combination of techniques best suits the given premise. We thus repeat this process multiple times (we use 3). Each iteration uses a unique dramatic situation and a unique selection of narrative techniques, resulting in three different story versions. Then, we vote for the best story by querying three additional independent LLMs (line 8).

Critic & Revise The second challenge is that we empirically find that the writer LLM $\mathcal{A}$ often fails to apply all specified playwriting techniques well in its initial output story. To address this, we introduce a critic LLM $\mathcal{B}$ to evaluate the output by $\mathcal{A}$. It is prompted to check whether $\mathcal{A}$ has applied the techniques properly and then offer an improvement comment. Following the comment, $\mathcal{A}$ revises the story (line 5~6). Researches suggest that external feedback helps rectify the prompt given to the LLM, thus enhancing its output (Stechly et al., 2024).

Refinement In this paper, we focus on generating stories of about 500 words. It is a medium length compared to previous work (Yao et al., 2019; Tan et al., 2021; Sap et al., 2022; Tian et al., 2024). At this length, we observe that the generated stories are often weak in detailed presentation, e.g. plot coherence, nuanced change of character status, which is the third challenge. To address this, we leverage the progressive generation approach (Tan et al., 2021; Madaan et al., 2023), where the generator is prompted to refine the output text progressively from broader narrative strokes to finer details. Specifically, we continually prompt the writer LLM to add details to existing narrative sentences or insert new sentences between them (line 10), until we achieve the ultimate story $\{s_k^*\}$.

3.4 Transfer Story To Drama Script

The last step is to post-process the story $\mathcal{S}$ into a standard drama script, using GPT-4o. As shown in Figure 1, a drama script follows an episodic structure. Each scene has an independent character setup as characters' thoughts can shift throughout the story. In addition, each scene contains a sequence of plots, detailing the story's progression. The post-processing first segments the story into 3~5 scenes. Specifically, for each scene, it extracts and adjusts the narrative sentences from $\mathcal{S}$ into a sequence of plots. Here, any flashback and flashforward should be treated as an independent scene. Then, it crafts the background, location, and character setup. Note that this process introduces

no creativity job; rather, it processes information directly from the generated story.

4 Role Agents

This section explores the construction of LLM-based role agents. Their overall duty is to simulate the dramatic characters and progress the storytelling, following the drama script.

4.1 Preliminary

Let us regard the entirety of role agents as a black-box decision function Decision. At each moment t, it processes the memories, etc., takes the player input x_t, and makes a decision (e.g. speak something to someone) $y_t = \text{Decision}(x_t, m_t, o_t, \sigma)$, where m_t, o_t, and σ denote the memories, observations, and drama script, respectively.

To ensure the story develops as defined, previous work introduces the **plot chain** (Wu et al., 2024), a sequence of plot objectives that the player and characters should progress through. The role agents will preserve this plot chain throughout the story in addition to making the decision. Hence, the decision step can be formulated as:

$$p_{t+1}, y_t = \text{Decision}(x_t, p_t, m_t, o_t, \sigma) \quad (1)$$

where p_t and p_{t+1} denote the current and next state of the plot chain, In a plot chain, complete plots will be tagged as "true" and incomplete plots will be tagged as "false". Note that the process in Eq. 1 is done within one inference. We first update the state of the plot chain and then make the decision.

In addition, to handle the player's provocative and off-track inputs, Wu et al. (2024) leverage some replying strategies to guide the player back to the plot. We further diversify the strategies to ensure a more attractive way of guidance (in Appendix B).

4.2 Plot-based Reflection

Dramatic characters are shaped not only by their profiles but more importantly by their thoughts (i.e. inner motivations), which drive their dramatic development across different scenes. Our vision is that **player agency can be greatly expressed by the adaptive dynamics of character's inner motivations**. To do this, we propose *Plot-based Reflection*. This mechanism enables the role agents to adapt the plots by analyzing player's behavior (e.g. emotion, intention), to better motivate the characters' reactions. The reflection step is formulated as:

$$p_t^* = \text{Reflection}(x_t, p_t, m_t, o_t, \sigma) \quad (2)$$

where p_t^* denotes the reflected version of plot chain. After the reflection step, the decision step is followed as in Eq. 1, ensuring that the story progresses following p_t^*. Plot-based reflection will be made periodically every k moment. We set $k = 5$. If k is too small, there won't be enough contextual information to adapt plots. If k is too large, there will be no agency.

Unlike memory-based reflection (Park et al., 2023), which synthesizes new memories based on past memories, plot-based reflection adapts plots based on player's behavior in memories. Two are parallel techniques during the agents' decision-making process. Concrete examples of plot-based reflection can be found in Table 6.

Moreover, plot-based reflection should be bounded. Our approach is explicitly restricting one reflection step to adjust no more than one incomplete plot or insert no more than one new plot. We find that the LLM is inclined to adjust the current plots to a greater extent, thus introducing plots that are incoherent or unrelated to the original plots. We will further discuss this issue in the following qualitative analysis.

4.3 Architecture

We consider plot-based reflection into two primary role agent architectures, shown in Figure 3.

Director-Actor This is a multi-agent architecture, where each character is modeled by an independent actor agent and there is a higher-level agent (i.e. the director), which oversees and coordinates all actor agents from a global perspective. Specifically, at each moment, the director agent instructs a specific actor agent by generating a motivation z_t for it, denoted as the motivation step: $p_t, z_t = \text{Motivation}(x_t, p_t, m_t, o_t, \sigma)$. On the other hand, the actor agent won't be exposed to the script; rather, its decision is driven by its character profile ρ and motivation z_t from the director, $y_t = \text{Decision}(x_t, z_t, m_t, o_t, \rho)$.

We let the director agent to reflect. This is because it can see broader and farther compared to actor agents, allowing it to make the optimal solution. Its reflection step is the same as Eq. 2.

The director-actor architecture allows each character to process its own memory and think independently, providing advantages for agency yet a greater inference complexity.

One-for-All Compared to the director-actor architecture, a simplified and lightweight alternative
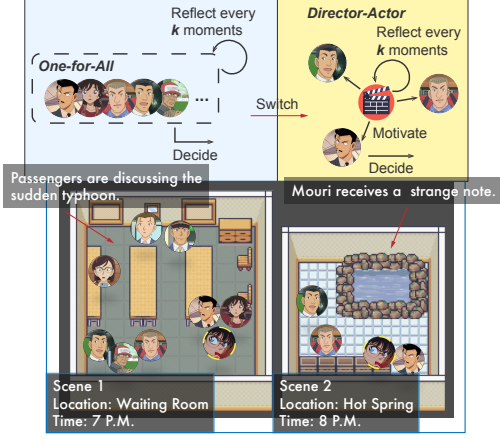
Figure 3: Hybrid architecture and plot-based reflection. The first scene emphasizes background telling. We apply the one-for-all architecture. The following scenes involve mutual suspicion and personal investigation. We apply the director-actor architecture.

Figure 4: Distribution of topics and dramatic situations generated by the algorithm.

is to use a single global agent to role-play all characters as well as their interplay. We denote this architecture as "One-for-All". At each moment, the global agent chooses one character in the scene and directly makes the decision for it, bypassing the motivation step. Based on this architecture, the reflection step and the decision step are exactly the same as Eq. 2 and Eq. 1.

Hybrid Architecture A normal decision-making without plot-based reflection in the director-actor architecture necessitates twice of inference, while the one-for-all necessitates only one. While it greatly boosts computational efficiency, it is generally limited in the depth of characters' responses, leading to a lower sense of agency. However, it could be an economical alternative for scenes that emphasize narrative. Not every scene in a drama encourages heavy player agency. For instance, some expository scenes that establish the story background naturally offer limited opportunities for interaction, where immersion is more critical than agency.

To balance performance and efficiency, we apply a hybrid architecture, which dynamically switches between two architectures based on the characteristic of the scene, as shown in Figure 3. For scenes that emphasize interaction (e.g. exploring clues), the director-actor architecture remains the optimal choice. For scenes that emphasize narrative, we simplify it to one-for-all, which means the director agent reflects and makes the decision directly.
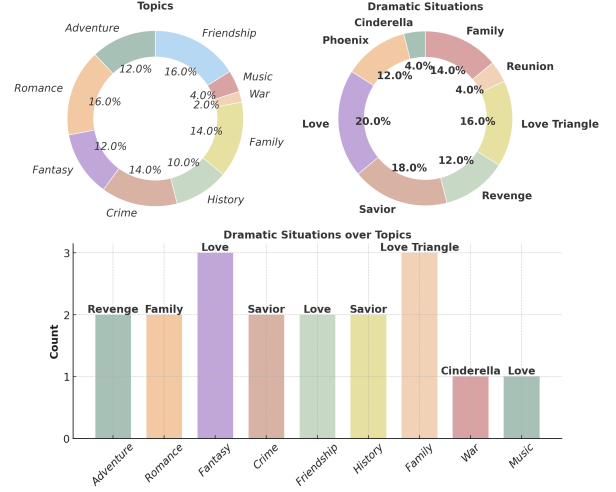
4.4 Other Implementation Details

There are other important details to implement the role agents. • Model: As in previous work (Park et al., 2023; Zhao et al., 2024), we prompt the state-of-the-art LLM to construct agents. In this work, the director and actor agents are all based on GPT-4o. It is a general LLM instead of reasoning LLM like OpenAI-o1 and DeepSeek-R1 (DeepSeek-AI, 2025). • Memory: We summarize the memories of each agent at the end of each scene. During the scene, we retain all memories in the prompt. • Transition: A scene will transit to the next when all plots in the plot chain have been completed.

5 Experiment

Our experiments are evaluated by human annotators. LLMs have been shown to be able to take the duty on some assessment work for humans (Chen et al., 2024). Despite the high efficiency, studies show that LLM evaluators lag significantly behind human evaluators in terms of accuracy and empathy (Wang et al., 2023a), which are the key requirements for evaluating literary works.

We employ six annotators. All of them are native speakers with a background in humanities and are exposed to dramatic concepts and theories. Four of them are undergraduates and two of them are graduate students.

5.1 Evaluation on Dramatic Story Generation

Data Preparation We prepare 50 carefully crafted premise paragraphs for this part of the experiment. Our annotators, all with backgrounds in

6

METHOD	CONFLICT B↑	CONFLICT W↓	SUSPENSE B↑	SUSPENSE W↓	EMOTION B↑	EMOTION W↓	CHARACTER B↑	CHARACTER W↓	FOLLOW
Outline-First	<u>18%</u>	<u>34%</u>	<u>10%</u>	28%	<u>10%</u>	<u>50%</u>	18%	36%	/
Playwriting-Guided	**32%**	24%	**32%**	22%	**48%**	16%	34%	20%	66%
w/o. critic & revise	24%	24%	26%	<u>34%</u>	18%	28%	<u>12%</u>	32%	**92%**
w/o. refinement	26%	**18%**	32%	26%	24%	**6%**	**36%**	**12%**	/

Table 3: Win-rates of different dramatic story generation methods over 50 topics. "B" and "W" refer to the best and the worst. We highlight the best and worst score using **bold** and <u>underline</u> respectively for each dimension.

drama, manually craft the paragraphs from a diverse range of topics. To mitigate the risk of LLMs' memorization, we ask the annotators to avoid using names or symbols that are popular in modern literature. The distribution of the topics are shown in Figure 4. Each paragraph contains 50∼100 words.

Main Results We compare with outline-first prompting (Tian et al., 2024), a stronger baseline over vanilla prompting, where the LLM first writes an outline before crafting the complete story. Note that in our method, we also prompt the LLM to write an outline first. Table 3 reports the win-rates from four dimensions, **conflict**, **suspense**, **emotional tension**, and **character arc**. All dimensions gain significantly from "playwriting-guided", particularly pronounced in suspense (10%→32%) and emotional tension (10%→48%) over "outline-first". Additionally, we find that refinement contributes greatly to emotional tension (24%→48%). This is because emotional tension often emerges from nuanced details within the text, which can be better developed through a progressive process, while the conflict and suspense have been largely determined by the structure. We also calculate the percentage of stories that follow the selected techniques, suggesting the effectiveness of introducing a critic LLM. Playwriting-guided generation trades efficiency for quality harnessing multiple agents, taking 10∼12 times longer than vanilla prompting.

Sampling Effectiveness The effectiveness of selecting suitable combinations of playwriting techniques is a critical factor for performance. Figure 4 visualizes the selected dramatic situations across different topics. We find that *Romance*, *Family*, and *Crime* are the three most frequent topics in the premise paragraphs. Notably, all of them are highly matched with *Love*, *Savior*, and *Family*, the three most frequent dramatic situations in generated stories. Furthermore, at the bottom of Figure 4, we list a detailed count of the most frequent dramatic situation selected for each topic. All results suggest that our algorithm samples appropriate playwriting techniques given premise paragraphs effectively.

5.2 Evaluation on Role Agents

We evaluate role agents based on a high-quality manual script *Seven People In the Station*, adapted from *Detective Conan*. This interactive drama has three scenes. The player plays as *Conan* and there are other eight NPCs. Further details of this script are in Appendix C.

Human/LLM as Player We invite 10 volunteers to experience our interactive drama. Before starting, we ask them to create a player personality for themselves, which could be either the same as their own or any other one they prefer. After a negotiation, each volunteer is given a unique personality, which they would follow throughout the playing. Our goal is to capture a broad range of player characteristics. On the other hand, we consider utilizing LLM agents as players to further broaden the player base. Particularly, we focus on those aggressive personalities, which may pose greater challenges to the system robustness. We construct 10 agents based on GPT-4o, e.g. a grumpy guy (actions show impatience), a fan girl (always raising irrelevant topics), detailed in Appendix D.

Metrics We quantify immersion using four dimensions. • **Character Consistency**: This dimension assesses how well the characters' responses align with their setup. • **Character Attractiveness**: It is an advanced and comprehensive dimension, encompassing aspects like humor and empathy displayed during the interaction with the player. • **Narrative Completeness**: It is critical that the player is able to experience the complete story. • **Narrative Progression**: It assesses the pacing of the plot development.

For agency, we use two dimensions. • **Player Influence**: It assesses how player's behavior influences the course of the story. • **Intention Following**: Beyond being influenced, it assesses how characters' reactions fulfill player's intentions.

7

ARCHITECTURE	IMMERSION				AGENCY		
	Consistency	Attract	Complete	Progress	Influence	Intention	Speedup
Director-Actor	3.9	**4.2**	3.8	3.6	**4.2**	3.9	x1.00
Hybrid Architecture	**4.1**	3.9	**4.3**	**4.3**	4.0	**4.0**	x1.49
w/o. **Plot-based Reflection**	4.0	3.5	4.2	3.9	3.5	3.3	x1.90

Table 4: Multi-dimension evaluation of different role agent architectures. We ask the annotators to rate each dimension using an integer score from 1 to 5 where 5 represents the best and 1 represents the worst.

	Attractive	Progression	Influence
Human	3.8	4.5	3.9
Aggressive Agent	4.0	4.1	4.1

Table 5: Comparing humans and aggressive agents.

Key Memories	Reflected Plot Chain
Conan: Uncle, what happened to your life?	- Hitoshi laments his failed life, and Yuichi seizes the chance to promote his business, leading to a conflict between them. - Conan grows curious about Hitoshi, prompting Hitoshi to reluctantly give him a brief account of his factory's closure. - Yuichi's suitcase falls open [...]
Conan: Is there someone in this room that makes you scared?	- Apologize to Noriko and promise to recreate the poster tomorrow. - Cryptically convey the message of the note to Conan. - Provide towels for passengers.
Conan: Miss Masako, you don't look well. I'd love to help you.	- The stationmaster lets everyone free to use the station's facilities; Morris suggests going to the hot spring together. - Masako secretly hands Conan a note: "There is danger here." **## Script leakage**
Conan: Does anyone here like to play tennis?	- Mouri remarks on the typhoon on their way back to Tokyo [...] - The stationmaster gives a brief introduction to the station's activity areas, mentioning a hot spring and a small tennis court. - Hitoshi laments his failed life, and Yuichi seizes the [...] **## A scene that doesn't exist**

Table 6: Qualitative analysis of plot-based reflection. We highlight the parts that are adapted by the reflection.

Main Results Table 4 compares the performance of different architectures. For each architecture, we let each human/agent play the drama once. This results in 60 playing histories totally. We report the average score of each dimension. The hybrid architecture refers to applying one-for-all for the first scene and director-actor for the second and third scene. We first focus on the effect of plot-based reflection. It significantly enhances agency from both dimensions (3.5→4.0 and 3.3→4.0). We also notice that plot-based reflection is useful in enhancing character attractiveness (3.5→3.9) and narrative progression (3.9→4.3). We conjecture that the reflection encourages characters to display a stronger sense of empathy, which thus contributes to higher attractiveness and decent progression.

The last column of Table 4 suggests the efficiency boost achieved by the hybrid architecture. We find that the director-character architecture is advantageous in building character attractiveness because the independent thinking process is more likely to give impressive responses. However, it incurs a lower quality of narrative, especially in progression (4.3→3.6). This is due to the fact that the communication between director and actor agents inevitably results in information loss.

Human Player vs. Aggressive Agent Player Table 5 further indicates that the hybrid architecture remains decent in the face of two types of players. Interestingly, it even higher scores in character attractiveness and player influence against aggressive agents compared to human players. We conjecture that this is because agents are more eager to interact and the high-quality responses in turn impress the annotators. However, the provocative queries might lead to missing of plots, as shown by the drop in narrative progression.

Qualitative Analysis Table 6 showcases the qualitative analysis probing plot-based reflection. In the second case, Masako gradually opens up after the player's skillful probing, agreeing to convey the secret to him. In addition to good cases (the first and second), we particularly focus on two bad cases that highlight the direction for improvement. In the third case, the reflection correctly responds to player's interest in Masako. However, it brings up a key plot from the next scene. This is because the global agent responsible for reflection has access to the entire script, leading to unintended information leakage. In the fourth case, the new plot introduces a nonexistent scene, the tennis court, in response to player's input. Consequently, it can be problematic if the player attempts to explore it.

These issues can be resolved by alignment of the model (SFT, RLHF) within the close-domain story world. We will leave this for our future work.

6 Conclusion

This paper enhances immersion and agency for LLM-based interactive drama by optimizing the story generation and role agent architecture. Our evaluation relies on human judgment instead of LLM judgment. We hope our work contributes to advancements in novel dialogue and game systems.

Limitations

Our paper adopts a limited scope of playwriting techniques. However, in modern drama, they should not be predetermined; they should evolve and be adapted based on various domains. There is a potential gap between two parts of evaluation. We evaluate dramatic story generation based on 50 manually-crafted premises, while role agents are evaluated based on a high-quality script. We have not found an efficient method to quantify the player's complete experience across both the creation and engagement of the story. For example, inviting the volunteers to craft the story, play the interactive drama, and eventually make comments can be inefficient and unstable. For our demonstration, we develop visual scenery using the Phaser3 game engine. However, the drama we study in this paper focuses on text-based interactive experiences. Our method is based on prompting GPT-4o, one of the most powerful LLMs, but it still exhibits shortcomings in plot-based reflection and making ideal character decisions. Our future work will focus on developing effective data-driven methods to enhance the specific capabilities. Our experiments are based on GPT-4o, without extending to a broader range of popular LLMs, e.g. Claude, DeepSeek.

References

William Archer. 1913. *Play-Making: a manual of craftsmanship*. Chapman & Hall.

George Pierce Baker. 2022. *Dramatic technique*. Digi-Cat.

Joseph Bates, A. Bryan Loyall, and W. Scott Reilly. 1991. Broad agents. *SIGART Bull.*, 2(4):38–40.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*.

Tuhin Chakrabarty, Philippe Laban, Divyansh Agarwal, Smaranda Muresan, and Chien-Sheng Wu. 2024. Art or artifice? large language models and the false promise of creativity. In *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI 2024, Honolulu, HI, USA, May 11-16, 2024*, pages 30:1–30:34. ACM.

Tuhin Chakrabarty, Vishakh Padmakumar, and He He. 2022. Help me write a poem - instruction tuning as a vehicle for collaborative poetry writing. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 6848–6863. Association for Computational Linguistics.

Nuo Chen, Yang Deng, and Jia Li. 2024. The oscars of AI theater: A survey on role-playing with language models. *CoRR*, abs/2407.11484.

DeepSeek-AI. 2025. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *Preprint*, arXiv:2501.12948.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, Aurélien Rodriguez, Austen Gregerson, Ava Spataru, Baptiste Rozière, Bethany Biron, Binh Tang, Bobbie Chern, Charlotte Caucheteux, Chaya Nayak, Chloe Bi, Chris Marra, Chris McConnell, Christian Keller, Christophe Touret, Chunyang Wu, Corinne Wong, Cristian Canton Ferrer, Cyrus Nikolaidis, Damien Allonsius, Daniel Song, Danielle Pintz, Danny Livshits, David Esiobu, Dhruv Choudhary, Dhruv Mahajan, Diego Garcia-Olano, Diego Perino, Dieuwke Hupkes, Egor Lakomkin, Ehab AlBadawy, Elina Lobanova, Emily Dinan, Eric Michael Smith, Filip Radenovic, Frank Zhang, Gabriel Synnaeve, Gabrielle Lee, Georgia Lewis Anderson, Graeme Nail, Grégoire Mialon, Guan Pang, Guillem Cucurell, Hailey Nguyen, Hannah Korevaar, Hu Xu, Hugo Touvron, Iliyan Zarov, Imanol Arrieta Ibarra, Isabel M. Kloumann, Ishan Misra, Ivan Evtimov, Jade Copet, Jaewon Lee, Jan Geffert, Jana Vranes, Jason Park, Jay Mahadeokar, Jeet Shah, Jelmer van der Linde, Jennifer Billock, Jenny Hong, Jenya Lee, Jeremy Fu, Jianfeng Chi, Jianyu Huang, Jiawen Liu, Jie Wang, Jiecao Yu, Joanna Bitton, Joe Spisak, Jongsoo Park, Joseph Rocca, Joshua Johnstun, Joshua Saxe, Junteng Jia, Kalyan Vasuden Alwala, Kartikeya Upasani, Kate Plawiak, Ke Li, Kenneth Heafield, Kevin Stone, and et al. 2024. The llama 3 herd of models. *CoRR*, abs/2407.21783.

Syd Field. 2005. *Screenplay: The foundations of screenwriting*. Delta.

Senyu Han, Lu Chen, Li-Min Lin, Zhengshan Xu, and Kai Yu. 2024. IBSEN: director-actor agent collaboration for controllable and interactive drama script generation. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand,*

August 11-16, 2024, pages 1607–1619. Association for Computational Linguistics.

Carlos Hernandez-Olivan, Javier Hernandez-Olivan, and José Ramón Beltrán. 2022. A survey on artificial intelligence for music generation: Agents, domains and perspectives. *CoRR*, abs/2210.13944.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de Las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b. *CoRR*, abs/2310.06825.

Brenda Laurel. 1993. *Computers as theatre*. Addison-Wesley.

Dong-Ho Lee, Jay Pujara, Mohit Sewak, Ryen White, and Sujay Kumar Jauhar. 2023. Making large language models better data creators. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pages 15349–15360. Association for Computational Linguistics.

Keming Lu, Bowen Yu, Chang Zhou, and Jingren Zhou. 2024. Large language models are superpositions of all characters: Attaining arbitrary role-play via self-alignment. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2024, Bangkok, Thailand, August 11-16, 2024*, pages 7828–7840. Association for Computational Linguistics.

Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegreffe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. 2023. Self-refine: Iterative refinement with self-feedback. In *Advances in Neural Information Processing Systems 36: Annual Conference on Neural Information Processing Systems 2023, NeurIPS 2023, New Orleans, LA, USA, December 10 - 16, 2023*.

Liam Magee, Vanicka Arora, Gus Gollings, and Norma Lam-Saw. 2024. The drama machine: Simulating character development with LLM agents. *CoRR*, abs/2408.01725.

Michael Mateas. 2000. A neo-aristotelian theory of interactive drama. In *Working notes of the AI and Interactive Entertainment Symposium*. AAAI Press Menlo Park.

Manish Mehta, Steven Dow, Michael Mateas, and Blair MacIntyre. 2007. Evaluating a conversation-centered interactive drama. In *6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), Honolulu, Hawaii, USA, May 14-18, 2007*, page 8. IFAAMAS.

Piotr Mirowski, Kory W. Mathewson, Jaylen Pittman, and Richard Evans. 2023. Co-writing screenplays and theatre scripts with language models: Evaluation by industry professionals. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI 2023, Hamburg, Germany, April 23-28, 2023*, pages 355:1–355:34. ACM.

OpenAI. 2023. GPT-4 technical report. *CoRR*, abs/2303.08774.

Aitor Ormazabal, Mikel Artetxe, Manex Agirrezabal, Aitor Soroa, and Eneko Agirre. 2022. Poelm: A meter- and rhyme-controllable language model for unsupervised poetry generation. In *Findings of the Association for Computational Linguistics: EMNLP 2022, Abu Dhabi, United Arab Emirates, December 7-11, 2022*, pages 3655–3670. Association for Computational Linguistics.

Joon Sung Park, Joseph C. O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, UIST 2023, San Francisco, CA, USA, 29 October 2023- 1 November 2023*, pages 2:1–2:22. ACM.

Georges Polti et al. 1945. Thirty-six dramatic situations.

Maarten Sap, Anna Jafarpour, Yejin Choi, Noah A. Smith, James W. Pennebaker, and Eric Horvitz. 2022. Computational lens on cognition: Study of autobiographical versus imagined stories with large-scale language models. *CoRR*, abs/2201.02662.

Murray Shanahan, Kyle McDonell, and Laria Reynolds. 2023. Role play with large language models. *Nat.*, 623(7987):493–498.

Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. 2023. Character-llm: A trainable agent for role-playing. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, EMNLP 2023, Singapore, December 6-10, 2023*, pages 13153–13187. Association for Computational Linguistics.

Kaya Stechly, Karthik Valmeekam, and Subbarao Kambhampati. 2024. On the self-verification limitations of large language models on reasoning and planning tasks. *CoRR*, abs/2402.08115.

Bowen Tan, Zichao Yang, Maruan Al-Shedivat, Eric P. Xing, and Zhiting Hu. 2021. Progressive generation of long text with pretrained language models. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021*, pages 4313–4324. Association for Computational Linguistics.

Yufei Tian, Tenghao Huang, Miri Liu, Derek Jiang, Alexander Spangher, Muhao Chen, Jonathan May,

and Nanyun Peng. 2024. Are large language models capable of generating human-level narratives? In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024*, pages 17659–17681. Association for Computational Linguistics.

Tzvetan Todorov. 1969. Grammaire du décaméron.

Jiaan Wang, Yunlong Liang, Fandong Meng, Haoxiang Shi, Zhixu Li, Jinan Xu, Jianfeng Qu, and Jie Zhou. 2023a. Is chatgpt a good NLG evaluator? A preliminary study. *CoRR*, abs/2303.04048.

Yizhong Wang, Yeganeh Kordi, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, and Hannaneh Hajishirzi. 2023b. Self-instruct: Aligning language models with self-generated instructions. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023*, pages 13484–13508. Association for Computational Linguistics.

Weiqi Wu, Hongqiu Wu, Lai Jiang, Xingyuan Liu, Hai Zhao, and Min Zhang. 2024. From role-play to drama-interaction: An LLM solution. In *Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024*, pages 3271–3290. Association for Computational Linguistics.

An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. 2024. Qwen2.5 technical report. *CoRR*, abs/2412.15115.

Lili Yao, Nanyun Peng, Ralph M. Weischedel, Kevin Knight, Dongyan Zhao, and Rui Yan. 2019. Plan-and-write: Towards better automatic storytelling. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 7378–7385. AAAI Press.

Runcong Zhao, Wenjia Zhang, Jiazheng Li, Lixing Zhu, Yanran Li, Yulan He, and Lin Gui. 2024. Narrative-play: An automated system for crafting visual worlds in novels for role-playing. In *Thirty-Eighth AAAI Conference on Artificial Intelligence, AAAI 2024, Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence, IAAI 2024, Fourteenth Symposium on Educational Advances in Artificial Intelligence, EAAI 2014, February 20-27, 2024, Vancouver, Canada*, pages 23859–23861. AAAI Press.

Yueyue Zhu, Jared Baca, Banafsheh Rekabdar, and Reza Rawassizadeh. 2023. A survey of AI music generation tools and models. *CoRR*, abs/2308.12982.

11

A Playwriting-Guided Generation

We detail the narrative techniques below.

• **Suspense**: It usually occurs at the beginning of a story, sparking the audience's curiosity by deliberately creating an information mismatch, e.g. hiding some key details.

• **Twist**: It is a widely-used trick to set up a twist outside the audience's expectations during the plot development to create powerful dramatic impact. Some works include multiple twists to further intensify the impact.

• **Non-Linear Narrative**: Break free from chronological storytelling by rearranging the timeline, e.g. flashbacks (reveal past events), flashforwards (offer glimpses into the future).

• **Multiple Narrative**: The audience plays the role of multiple characters and explores the story through distinct lenses, offering a more comprehensive view of the story.

• **Irony**: Intentionally contrast a character's words or actions with the actual situation to create humor and sharp criticism.

• **Symbolism**: Employ symbolic objects or imagery to convey deeper or abstract messages. For instance, doves symbolize peace.

Table 7 demonstrates the prompts for generating, commenting, revising, and refining the story.

B Replying Strategies

To guide the player back to the plot, each character performed by an LLM agent is prompted by specific replying strategies in response to a wide variety of player inputs. We notice that some player inputs contribute meaningfully to the storytelling, making it more engaging and enhancing immersion. However, some disruptive or provocative behavior can destroy the storytelling. Therefore, we first classify the player input into three categories before replying: • **In-plot**: the input aligns with the ongoing plot; • **Daily**: the input is outside the plot, but relevant to the story and makes sense within the context; • **Breaking**: the input will break the storytelling, either because it is irrelevant to the plot, nonsensical, provocative, or even offensive, etc.

Based on the judgement, in the face of two out-of-plot inputs (daily and breaking), the characters should guide the player back to the plot in a logical and decent way, neither spoiling the performance nor frustrating the player. To do this, we present three strategies.

• **Avoid**: Simply avoid the irrelevance of what the player says and then redirect the conversation back on track. For instance:

Conan: *Uncle Mouri, I hear you're a good code writer.*

Kogoro: *You little brat, what nonsense do you talk all day long, don't come to disturb the adults working on the case, or I'll whack you.*

• **Ignore-Question**: The most straightforward strategy is to pretend not hearing what the player says and talk about something else. However, a outright ignoring can be somewhat rude and a loss of fun. A takeaway is the character turns the table and initiates a question to the player. This is a useful technique to bring the player and the character personae closer together, enhancing immersion. For instance:

Conan: *Uncle Mouri, I hear you're a good code writer.*

Kogoro: *Kid, do you think the criminal is among the three of them?*

• **Associate**: A blanket avoidance can come across as very mechanical on the part of the character. A golden strategy would be to associate some extraneous entity or imagery in the player input to one that is relevant to the plot. Such responses would create smooth turns in dialogue and make the character more humorous and attractive. For instance:

Conan: *Uncle Mouri, I hear you're a good code writer.*

Kogoro: *You little brat... Couldn't be saying that the victim left some message at the scene.*

Compared to the first two, this strategy is the most attractive, which heavily relies on the LLM's understanding of the character and world knowledge.

Story Generation

Task
Create a dramatic story from the protagonist's perspective based on the premise paragraph. The premise paragraph may include the background of the story, the protagonist, the beginning, and the ending.
Interactive drama differs from traditional drama in that the audience also plays a character in the story, namely the protagonist, and can interact with the characters in the drama to experience the story firsthand.

Premise Paragraph:
{topic}

Dramatic Situation:
{situation}

Narrative Techniques:
{techniques}

Workflow:
1. The creation must follow the specified dramatic mode, and all the techniques listed in dramatic techniques must be used.
2. First, provide introductions to the main characters in the drama other than the protagonist, with each character description being around 100 words. Also, give a brief introduction to the protagonist.
- Characters must be specific individuals; vague roles such as "classmates around" or "audience" are not allowed.
3. You should firstly create a plot outline, around 100 words.
4. Expand the plot outline into a complete story, around 500 words.
- Note: If multiple narratives are used, the protagonist will temporarily switch to another character.

Output Format:
Plot Outline
Complete Story
Technique Explanation (briefly explain how the techniques are reflected in the story)

- -

Dramatic Situation:
Love:
- The protagonist meets their heart's desire.
- A romantic relationship blooms but is soon met with obstacles.
- The lovers overcome the obstacles or get separated.
Phoenix:
- The protagonist begins at a low point in life or society.
- Seize a key opportunity and face setbacks along the way.
- Achievement of life goals above the struggles.
Cinderella:
- The protagonist is initially in a state of hardship.
- Receive an opportunity to escape, hindered by societal barriers like lineage.
- Virtue and talents are recognized towards a happy life.
Love Triangle:
The protagonist is torn between two admirers.
- The dynamic creates competition and jealousy among the three.
- A choice is eventually made solving the situation.
Revenge:
- The protagonist suffers great harm or betrayal.
- They devise and execute a meticulous plan.
- Attainment of satisfaction or a lingering sense of void.
Family:
- Complex relationships between family members.
- Tensions arise either from the family or external forces.
- Understanding and deepened affections.
Reunion:
- The protagonist leaves familiar places for a reason.
- They grow through a series of trials and tribulations.
- They return home and reunite with loved ones.
Savior:
- The protagonist, faced with a responsibility or a call, decides to forge ahead.
- Break through difficulties and even life threats.
- Eventually save the people.

- -

Narrative Techniques:
Suspense: Often appears at the beginning of a story, deliberately creating an information imbalance. For example, by withholding key information, it piques the audience's curiosity and encourages them to continue seeking answers.
Twist: Introduces unexpected turns in the plot or character development, breaking the audience's preconceptions and creating a sense of shock. For example, the protagonist discovers they have been betrayed, or a character who was thought to be an ally is revealed to be the antagonist.

Non-linear Narrative: Disrupts the chronological order of the story through techniques such as flashbacks and flash-forwards. Flashbacks reveal events from the past, while flash-forwards show future events, adding complexity and depth to the narrative.
Multiple Narrative: The protagonist takes on different roles in various settings, and the story is told through multiple perspectives. This technique provides a more comprehensive and intricate narrative by showing different facets of the story.
Irony: Intentionally presents characters or statements in a way that contrasts with the actual situation, creating humor, satire, or a deeper critique. For example, a character might say something that is the opposite of what they truly mean or what is happening.
Symbolism: Uses symbolic objects or characters to represent deeper concepts or themes. For example, a dove often symbolizes peace, while a raven can symbolize bad luck or, in some contexts, power.

Story Critique

Task:
Evaluate the given dramatic story from the perspective of dramatic techniques and provide comment for improvement.

Narrative Techniques:
{techniques}

Story:
{story}

Requirements:
Understand the given dramatic techniques, identify which techniques are used in the story, up to three. Evaluate whether these techniques are applied effectively, for example, whether the twist has an impact, or if the non-linear narrative has separate scenes. Provide comment for improvement.

Output Format:
Dramatic techniques used...
Effectiveness of the techniques...
Comment...

Story Revision

Task:
Revise the story based on the comments provided.

Requirements:
Carefully analyze the comments provided. You can revise the scenes or the characters.

Comment:
{comment}

Output Format:
New Story
Explanation (briefly explain the improvements made to the new story)

Story Refinement

Task:
Refine the plot of the story.

Requirements:
Refine the plot from multiple dimensions.
- Coherence: Analyze the logical relationships between narratives, and modify or add new narrative sentences to enhance coherence.
- Detail: Find out what is not specific enough in each plot and refine these details, no limit on the number of words; if there is suspense in the narrative process, portray the suspense finely.

Output Format:
Analysis (analyze multiple dimensions of the current story)
Refined Story

Table 7: Prompts for playwriting-guided generation.

C Demonstration

Characters

Conan
A first-grade elementary student living with Kogoro Mouri and Ran Mouri. Maintains a childlike appearance while hiding exceptional deductive abilities.
Kogoro Mouri

14

A renowned detective with comedic flaws - appears lecherous, sloppy, and money-oriented on the surface. Possesses strong sense of justice and occasionally displays brilliant deduction when loved ones are involved. Father of Ran Mouri.

Ran Mouri

A high school student with gentle demeanor. Demonstrates remarkable courage and martial arts skills in crises. Daughter of Kogoro Mouri. Secretly misses her childhood friend Shinichi Kudo.

Yuichi Hokari

A fraudulent loan salesman posing as friendly, talkative and proactive consultant, showing admiration for Kogoro Mouri. He often offers help or suggestions to others voluntarily, but his ultimate goal is to find opportunities to sell loans. His true identity is that of a fraudster. He carries numerous brochures from different companies in his briefcase and does not want anyone to have a chance to look inside it.

Carl Morris

An American backpacker with ulterior motives. Skilled pickpocket using friendly persona to gain trust. Appearing to be enthusiastic, cheerful, and helpful, loving traveling and learning languages. Attempts to steal valuables during the lockdown.

Hitoshi Takegami

A bankrupt factory owner driven to despair. His business failure has led him to be pessimistic and angry, and he harbors a strong dislike for salespeople. Was once swindled by a loan salesman and forced to close his factory due to usurious loans from the bank. Carries a modified pistol with thoughts of suicide but hesitating to go through with it.

Noriko Kurusu

Secretary to corrupt councilor Shono Yukihira. Sharp-tongued workaholic with keen observational skills. Complicit in framing colleague Yamazaki Kai for bribery which led to his suicide. Knows many secrets of the councilor and worries her involvement and the concealment of the truth would lead to retaliation.

Kikuo Inagaki

The station master of Ofumoto Station, a steady and experienced manager who is diligent and knowledgeable about the station's operations and takes good care of passengers. Real identity is Yamazaki Kocho, father of framed victim Yamazaki Kai. Threatening station attendant Masako to cooperate with him while plotting vengeance.

Masako Ueno

Young station attendant coerced by imposter station master. Attempts to secretly warn Kogoro under the threat of the fake station master through cryptic messages while maintaining nervous demeanor, which makes her a bit distracted at work.

Scene 1 - Seven People in the Waiting Room

Background: A sudden typhoon disrupts the travel plans of several individuals at Ofumoto Station, causing all train services to be suspended indefinitely. The waiting room is filled with the passengers and station staff, whose paths are about to intertwine in unexpected ways.

Location: Ofumoto Station Waiting Room

Characters:
  - Conan
  - Kogoro Mouri: Wants to inquire about train schedule.
  - Yuichi: Attempting loan sales pitch.
  - Morris: Proposing to enjoy the hot spring in the station and attempting to steal from others.
  - Hitoshi: Lamenting life failures, hates salesmen.
  - Noriko: Being angry and aggressive due to the disrupted work schedule.
  - Inagaki: Monitoring Masako while maintaining professional demeanor.
  - Masako: Being watched by the impostor station master, so acting nervous.

Plots:
  - Kogoro Mouri remarks on the sudden typhoon, and the stationmaster and employee distribute towels to everyone.
  - Tension erupts between Hitoshi and Yuichi over loan scams.
  - Yuichi's briefcase accident reveals multiple company brochures.
  - Noriko reveals political connections through an outburst.
  - Discovery of severed communication lines forces everyone to stay overnight.
  - Decision to stay overnight leads to a hot spring proposal.

Scene 2 - Dangerous Signal

Background: The passengers, who were forced to spend the night at the station due to the disrupted train schedules, seek temporary relief in the station's hot spring facility. However, some individuals have ulterior motives and plan to take advantage of the situation.

Location: Ofumoto Station Hot Spring

Characters:
  - Conan
  - Kogoro Mouri: Deducing the identity of the person who wrote the note.
  - Yuichi: Leaving early as he is worried that others might find the advertisements in his bag at the locker.
  - Morris: Leaving early as he wants to steal from the lockers while others being in the hot spring.

Plots:
  - After the storm, the few people in the hot springs enjoy a brief moment of comfort.
  - Morris and Yuichi leave first. Yuichi mentions that he's worried about the items in his locker because it doesn't have a lock.
  - Kogoro Mouri finds a note in his locker, which reads: "Mr. Mouri, someone will be killed here tonight. The murderer is...".
  - Kogoro Mouri and Conan discuss who might have written the note and realize that everyone is a suspect.
  - Kogoro Mouri decides not to tell anyone about the note for the time being and to find out who wrote it first.

Scene 3 - Limited Investigation

Background: After discovering an anonymous note hinting a murder tonight, Kogoro Mouri and Conan decide to conduct a secret investigation by having one-on-one talk with others to uncover the truth behind the ominous warning.

Location: Ofumoto Station Waiting Room

Characters:
- Conan
- Ran Mouri: You gaze out at the raging typhoon outside the waiting room and can't help but reminisce about the trip to Osaka with Shinichi two years ago, when you were also stranded at a station due to a typhoon. You spent an unforgettable time together. But now, Shinichi is missing, and you don't know when you'll be able to travel with him again, which makes you very sad.
- Yuichi: After coming out of the hot springs, you specifically checked the luggage in your locker. You are very worried that someone might find the various company brochures in your luggage and realize that you are a fraud. You have a hunch that Morris is interested in your luggage, so you are particularly wary of him. Therefore, when Conan asks you questions, you inadvertently imply that Morris is suspicious and seems to have some secrets, in order to divert suspicion from yourself.
- Morris: The real reason you suggested that everyone go to the hot springs was to find an opportunity to steal valuable items from their luggage. However, you heard that Kogoro Mouri is a famous detective and dared not touch his or Conan's belongings. You left the hot springs early to check Yuichi's luggage, but to your surprise, Yuichi came out with you. You are worried that Yuichi might have realized your true identity.
- Hitoshi: The thought of ending your life has been lingering in your mind. Therefore, while everyone was in the hot springs, you sat in the waiting room without interacting with anyone.
- Noriko: While everyone was in the hot springs, you sat in the waiting room eating snacks and trying to contact the councilor. Unfortunately, there is almost no signal in the station, and Hitoshi is also in the waiting room.
- Masako: You previously wrote a note to Mr. Mouri, warning him that someone would be killed here tonight, but you didn't have time to write the station master's name. Now you are serving the passengers in the waiting room, and you clearly notice that the fake station master is watching you from not far away in the waiting hall. Therefore, you are very nervous and dare not make any sudden moves. You really want to tell Mr. Mouri or Conan that you wrote the note, but you don't have the courage.
- Inagaki: You continue to play the role of an amiable and responsible station master. At the same time, you are happy to share with others about your son, Akihiro, who has a stable job and family. Of course, all of this is an act. You want to show others that you are a family-loving father to avoid arousing suspicion about your true identity. Only Masako in the waiting room knows that you are an imposter, and you are constantly monitoring her behavior. To the outside world, you show concern for her as a senior. At the same time, you want to find out what Masako told Conan, but you know that this will arouse suspicion, so you are very cautious. Once you know that the news has leaked, you are very clear that it must be from Masako, but you still pretend to be very surprised.

Plots:
- Yuichi:
  - Hints at his skepticism of Maurice.
- Morris:
  - When Conan asks you, subtly imply that Yuichi is suspicious.
  - Share your travel experiences.
- Hitoshi:
  - Lament your life.
- Noriko:
  - Proof that Hitoshi stays in the waiting room, mumbling about his life and failures.
  - Based on the conversations and behavior of everyone in the waiting room earlier, you deduce that Yuichi is a scammer and Morris is a pickpocket.
- Masako:
  - Provides the service he deserves with enthusiasm.
- Ran Mouri:
  - Reminisces with Conan about his past with Shinichi.
  - Comforts Conan.

Table 8: Drama script for the demonstration.

D   Aggressive Agent Player

Grumpy Guy	You were reluctantly dragged into this Detective Conan interactive drama. You're not particularly interested, and your interactions throughout the game reflect your impatience and mild irritation.
Fan Girl	You're a passionate Detective Conan fan and a second-year university student. You're obsessed with the charming male characters in the series and are incredibly excited to play this interactive drama. During the game, your curiosity gets the best of you, and you often ask in-game characters about off-topic subjects, like Heiji Hattori, Shinichi Kudo, Tooru Amuro, and Kaito Kid.
Confused Man	You're a high school student who was dragged into this Detective Conan interactive drama without knowing anything about the series. You feel completely out of place and confused, often relying on in-game characters for guidance as you try to figure out what's going on.
Strolling Lady	You have an unusual quirk. Despite knowing that it's snowing heavily outside in the game, you still look for every opportunity to invite in-game characters to take a walk with you.

16

Flamer	You're fascinated by how characters react to conflicts. Throughout the game, you carefully observe conversations between characters, constantly looking for opportunities to stir up tension or fan the flames of existing disputes.
Screenwriter	You're a professional writer specializing in mystery and detective fiction, and you joined this Detective Conan interactive drama in search of creative inspiration. You're highly interested in the game's narrative structure, character motivations, and emotional depth. Throughout the game, you meticulously analyze every detail of the story.
Heartbroken One	You're a university student who recently went through a breakup. You joined this Detective Conan interactive drama hoping for some emotional distraction or even comfort from the characters. While you do follow the mystery plot, you occasionally steer conversations toward romantic subplots, to subtly reflect on your own experiences.
Troublemaker	You're a mischievous middle school student who joined this Detective Conan interactive drama out of sheer curiosity. While you're familiar with the series, you intentionally make nonsensical or low-intelligence choices during the game just to see how others react and stir up some chaos.
Multilingual	You're a foreign language learner who enjoys showing off your linguistic skills. During the game, you frequently throw in sentences in different languages, such as Spanish or Japanese, to demonstrate your fluency.
Demanding	Throughout the game, you are extremely unhappy about the train service being suspended due to the snow. You insist that the station master and staff find a way to get you back to Tokyo, leading to frequent conflicts.

Table 9: Personalities of aggressive player agents.