

# AdaMARP: An Adaptive Multi-Agent Interaction Framework for General Immersive Role-Playing

Zhenhua Xu<sup>1†</sup>, Dongsheng Chen<sup>2†</sup>  
 Shuo Wang<sup>2</sup>, Jian Li<sup>2</sup>, Chengjie Wang<sup>2</sup>, Meng Han<sup>2</sup>, Yabiao Wang<sup>2</sup>

<sup>1</sup>Zhejiang University, <sup>2</sup>Tencent YouTu Lab

LLM role-playing seeks to portray arbitrary characters in interactive narratives, yet existing systems often lack immersion and adaptability. They typically under-model *dynamic environment information* and assume a largely static scene/cast, offering limited support for multi-character orchestration, scene transitions, and on-the-fly character introduction. We propose an adaptive multi-agent interaction framework dubbed **AdaMARP**, which featuring an immersive message format that interleaves [Thought], (Action), <Environment>, and Speech, and an explicit **Scene Manager** that controls role-playing via discrete actions (`init_scene`, `pick_speaker`, `switch_scene`, `add_role`, `end`) with rationales. To train these abilities, we construct **AdaRPSet** for the Actor Model and **AdaSMS** for supervising orchestration decisions, and introduce **AdaptiveBench** for trajectory-level evaluation. Experiments across multiple backbones and scales show consistent gains: AdaRPSet improves character consistency, environment grounding, and narrative coherence—an 8B actor outperforming several commercial LLMs, while AdaSMS enables smoother scene transitions and more natural role introductions, surpassing Claude Sonnet 4.5 with only 14B LLMs.

- 📅 **Date:** Jan. 16, 2026
- ✉️ **Correspondence:** [mhan@zju.edu.cn](mailto:mhan@zju.edu.cn)
- 🌐 **Project Leader:** [caseywang@tencent.com](mailto:caseywang@tencent.com)
- 🌐 **Project Homepage:** <https://xuzhenhua55.github.io/AdaMARP>

## 1 Introduction

Recent advances in large language models (LLMs) have substantially improved their general-purpose capabilities across many language tasks [Tan et al., 2025, Zhang et al., 2025, Zhuang et al., 2025, Zhang et al., 2025, Xu et al., 2025, Kong et al., 2025, Xu et al., 2025, Li et al., 2025, Xu et al., 2025]. One prominent application is *LLM role-playing*—prompting an LLM to portray a character and interact with a user in an ongoing narrative—which has gained rapid popularity [Shanahan et al., 2023, Tseng et al., 2024, Chen et al., 2024] and spawned a growing ecosystem of commercial products<sup>1</sup>. This motivates research on **general role-playing**: enabling an LLM to convincingly portray *arbitrary* user-defined or fictional characters [Wang et al., 2025], rather than a small pre-defined roster [Yu et al., 2024, Yang et al., 2025].

Despite this progress, existing general role-playing models often fall short in **immersion**. A large body of prior work focuses primarily on generating verbal Speech alone [Shao et al., 2023, Chen et al., 2023, Li et al., 2023, Wang et al., 2024, Zhou et al., 2024, Lu et al., 2024, Zhou et al., 2024]. Subsequent efforts incorporate Action [Tu et al., 2024, He et al., 2025] or additionally model Thought (e.g., CoSER) [Wang et al., 2025]. However, a crucial signal remains under-modeled: **dynamic environment information**. In narratives, environment is not mere decoration—it shapes atmosphere and causality, linking actions, world changes, and subsequent dialogue. This motivates **RQ1**: *How can we design a more immersive message representation that tightly couples character behavior with evolving environments?*

Most role-playing systems assume a **static interaction structure**: a fixed scene and a fixed character interacting with

<sup>†</sup> Equal contribution.

<sup>1</sup><https://character.ai/>  
<https://reprika.ai/>  
<https://spicychat.ai/>

the user. Even when multiple characters are involved [Chen et al., 2023, Wang et al., 2025], existing frameworks seldom provide a *systematic* way to choose the next speaker with a brief rationale, and rarely support narrative-level dynamics such as scene transitions or on-the-fly character introduction as the plot evolves. These limitations motivate **RQ2**: *How can we design a role-playing framework that supports dynamic multi-character orchestration, scene transitions, and on-the-fly character introduction?*

To address RQ1, we propose an **immersive messaging configuration** that explicitly interleaves four elements—[Thought], (Action), <Environment>, and Speech—which may appear flexibly within a single turn. This formulation enhances situational grounding and facilitates narratives in which environmental states and their dynamics actively participate in the interaction loop.

To address RQ2, we formalize the notion of **adaptive role-playing**, where the system (i) dynamically selects the most appropriate next speaker among multiple roles (including the user), (ii) switches scenes as the plot progresses, and (iii) introduces new characters when required by the narrative. Based on this formulation, we present **AdaMARP** (An Adaptive Multi-Agent Interaction Framework for General Immersive Role-Playing), which models role-playing as the interaction of three agents: an Actor Model that portrays all non-user roles, a User Model (simulated by an LLM or replaced by a real user), and a Scene Manager that performs high-level control. The Scene Manager operates over a discrete action space (`init_scene`, `pick_speaker`, `switch_scene`, `add_role`, `end`) and outputs explicit rationales to guide subsequent generation; in particular, it always begins an episode with a single `init_scene` action that establishes the initial scene context.

To train models that can follow this adaptive framework, we construct two datasets. First, we build **AdaRPSet** (Adaptive Role-Playing Dataset) for training the Actor Model, consisting of (i) **AdaRPSet-Extracted**, extracted from narrative books with LLM-assisted scene extraction and profile synthesis, and (ii) **AdaRPSet-Synthesis**, an LLM-synthesized corpus explicitly covering dynamic phenomena such as scene switching and role addition across 20 topics. Second, to improve orchestration quality, we construct **AdaSMSet** (Adaptive Scene Manager Dataset) based on the synthesized trajectories to supervise the Scene Manager. This dataset includes inserted `pick_speaker` decisions and their rationales.

Finally, to evaluate the adaptive role-playing abilities of the model beyond sentence-level metrics, we propose **AdaptiveBench**, a simulation-based benchmark that generates full dialogue trajectories under Scene-Manager control and evaluates performance at the trajectory level. We assess the Actor Model with a rubric covering character consistency, environmental grounding, interpersonal interaction, narrative progression, and instruction compliance, and evaluate the Scene Manager in terms of scene understanding, speaker discipline, role-introduction judgment, and overall quality.

Experiments across multiple backbones and model scales demonstrate the effectiveness of our data and supervision strategy. **AdaRPSet** consistently improves the Actor Model in AdaMARP, yielding stronger character consistency, richer environment-grounded narration, and more coherent progression under dynamic scene and cast changes. Notably, an 8B Actor Model trained with AdaRPSet outperforms several proprietary role-playing LLMs (e.g., GPT-4o-mini, Gemini-2.5-Pro, and Doubao-1.5-Pro-Character) in our trajectory-level evaluation. **AdaSMSet** further enhances the Scene Manager’s orchestration quality, allowing it to surpass Claude Sonnet 4.5 at the 14B scale. This is reflected in more appropriate scene transitions and better-justified role introductions.

Our main contributions are summarized as follows:

- We propose **AdaMARP**, an adaptive framework that integrates an environment-aware message format (interleaving thought, action, environment, and speech) with a discrete-action Scene Manager to enable dynamic multi-character orchestration, supporting both scene transitions and on-the-fly role additions.
- We construct **AdaRPSet** and **AdaSMSet**, two large-scale datasets designed to supervise immersive role portrayal and high-level narrative control (e.g., scene transitions, role additions), respectively.
- We introduce **AdaptiveBench** for trajectory-level evaluation **in adaptive role-playing**. Experiments demonstrate that our 8B Actor Model outperforms GPT-4o-mini, while our 14B Scene Manager surpasses Claude Sonnet 4.5.

Dataset	Character							Conversations					Message Components						
	I&A	P&P	SS	AIA	SHC	PHA	#Char	Init.	Scene	Multi-Char	Interloc.	Rel.	#Dial.	Avg. Turn	Speech	Action	Thought	Env.	Open
ChatHaruhi [Li et al., 2023]	✓	✗	✗	✗	✗	✗	32	✗	✗	✗	✗	✗	54,726	Fixed 2	✓	✗	✗	✗	✓
CharacterLLM [Shao et al., 2023]	✓	See Appendix A		✗	✓	✗	9	✓	✗	✓	✓	✓	14,300	13.2	✓	✗	●	✗	✓
HPD [Chen et al., 2023]	✓	✓	✗	✓	✗	✗	113	✓	✓	✓	✓	✓	1,042	13.8	✓	✗	✗	✗	✓
RoleLLM [Wang et al., 2024]	✓	See Appendix A		✗	✗	✗	100	✗	✗	✗	✗	✗	168,093	Fixed 2	✓	✗	✗	✗	✓
CharacterGLM [Zhou et al., 2024]	✓	●	✗	●	✓	✓	250	✗	✗	✗	✓	✓	1,034	15.8	✓	✗	✗	✗	✓
DITTO [Lu et al., 2024]	✓	✗	✗	✓	✓	✓	4,002	✗	✗	✗	✗	✗	7,186	5.10	✓	✗	✗	✗	✓
ROLEPERSONALITY [Ran et al., 2024]	✓	See Appendix A		✗	✗	✗	46	✗	✗	✗	✗	✗	32,089/32,767	Fixed 2 or 5	✓	✗	✗	✗	✓
SimChat [Yang et al., 2025]	✓	✓	✓	✗	✗	✗	68	✓	✗	✓	✓	✗	13,971	10.3	✓	✗	●	✗	✓
BeyondDialogue [Yu et al., 2025]	✓	✓	✓	✗	✓	✗	331	✓	✓	✓	✓	✓	3,552	6.54	✓	✗	✗	✗	✓
Crab [He et al., 2025]	✓	✓	✓	✓	✗	✗	18,424	✗	✗	✓	✓	✓	41,631	4.96	✓	✓	✗	✗	✓
TAILORGEN [Gao et al., 2025]	✓	See Appendix A		✗	✗	✗	5	✗	✗	✗	✗	✗	7,671	Fixed 2*	✓	✗	✗	✗	✗
CoSER [Wang et al., 2025]	✓	✓	✓	✓	✓	✓	17,966	✓	✓	✓	✓	✓	29,798	13.20	✓	✓	✓	●	✓
<b>Ours</b>	✓	✓	✓	✓	✓	✓	29,335	✓	✓	✓	✓	✓	22,425	20.08	✓	✓	✓	✓	✓

**Table 1.** Comparison of general role-playing datasets. Profile subfields are abbreviated as: I&A (Identity & Appearance), P&P (Personality & Psychology), SS (Speaking Style), AIA (Abilities, Interests & Achievements), SHC (Social & Historical Context), and PHA (Personal History Arc). #Char denotes the number of distinct characters in the dataset. Init. Scene indicates whether dialogues are provided with an explicit initial scene or setting. Interloc. indicates whether the user (or dialogue partner) is assigned a specific profile. Rel. denotes whether explicit relationships among characters are modeled or annotated in dialogues. #Dial. refers to the total number of dialogues. Env. is short for Environment, indicating whether environmental descriptions are included. Open indicates whether the dataset is open-sourced. ● and \* denote partial/implicit support or values inferred from the reported pipeline; detailed clarifications on symbol interpretations and dataset releases are provided in Appendix A.

## 2 Related Work

**Character-Specific Role-Playing.** Many approaches specialize an LLM to one or a few fixed characters, achieving strong fidelity but limited generalization to new roles (e.g., CharacterLLM [Shao et al., 2023], Neeko [Yu et al., 2024], HyCoRA [Yang et al., 2025]). Consequently, they are ill-suited for open-ended scenarios where the model must adopt arbitrary user-defined personas without per-character retraining.

**General Role-Playing via Data Construction.** General role-playing requires models to portray arbitrary user-defined or fictional characters. A prevalent line of work constructs large-scale multi-character corpora to improve persona coverage and avoid overfitting to a fixed roster, e.g., ChatHaruhi [Li et al., 2023], DITTO [Chen et al., 2023], CharacterGLM [Zhou et al., 2024], ROLEPERSONALITY [Ran et al., 2024], SimChat [Yang et al., 2025], BeyondDialogue [Yu et al., 2025], Crab [He et al., 2025], TAILORGEN [Gao et al., 2025], and CoSER [Wang et al., 2025]. While these efforts improve dialogue quality (sometimes incorporating thought/action), most do not explicitly model dynamic environments as first-class signals, nor provide systematic supervision for scene transitions and cast changes. For a systematic comparison, we summarize the key differences between existing datasets and ours in Table 1.

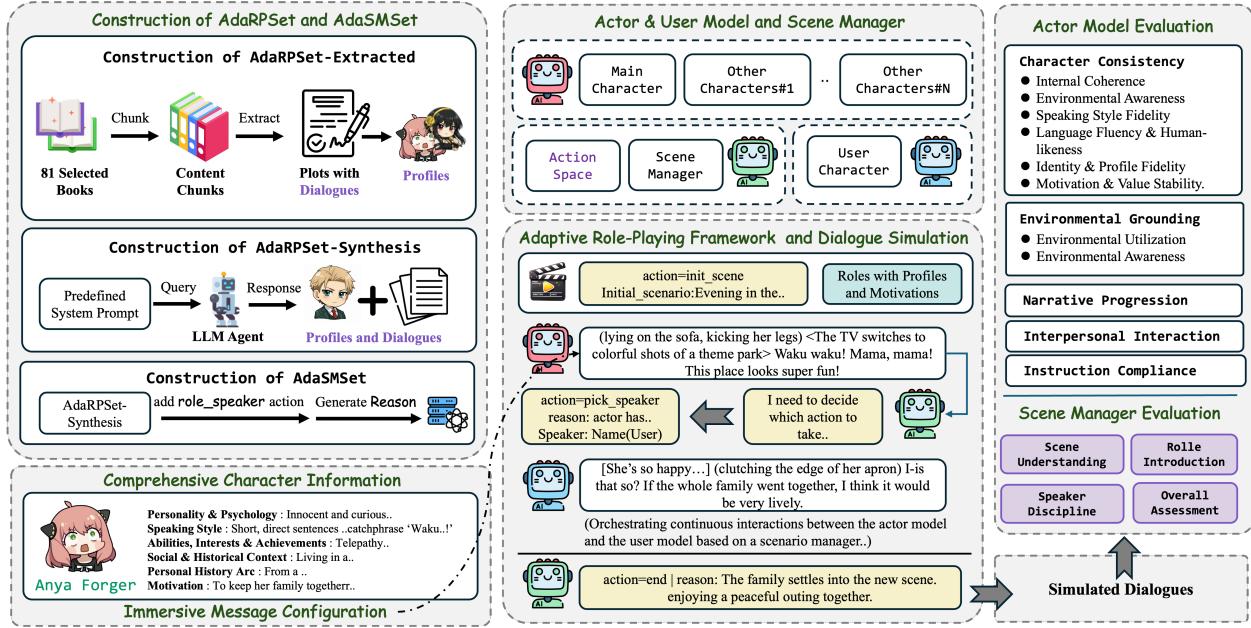
**General Role-Playing via Training Strategies.** Another line improves role-playing with specialized objectives or optimization, such as persona-aware contrastive learning (PCL) [Ji et al., 2025], metacognition-driven training (R-CHAR) [Qin et al., 2025], and RL variants that address reward noise or design implicit rewards (CPO [Ye et al., 2025], CogDual [Liu et al., 2025]). Our work targets general role-playing with supervised fine-tuning, so we mainly compare against data-construction approaches; training-strategy methods are largely complementary.

## 3 Method

### 3.1 Design Principles of AdaMARP

#### 3.1.1 Comprehensive Character Information

To support nuanced and consistent role-playing, AdaMARP adopts a structured profile for the main character that covers seven complementary dimensions: (I) Identity & Appearance, (II) Personality & Psychology, (III) Speaking Style, (IV) Abilities, Interests & Achievements, (V) Social & Historical Context, (VI) Personal History Arc, and (VII) Relationships. Together, these dimensions provide both intrinsic traits and narrative/social grounding, enabling coherent behavior, stable interaction patterns, and evolution-aware characterization across scene changes and long trajectories. Detailed specifications are provided in Appendix B.



**Figure 1.** Overall framework of AdaMARP. The left part illustrates the construction of AdaRPSet and AdaSMSet and the components of a comprehensive role profile. The upper middle part shows the roles played by the three agents, while the lower middle part depicts the dialogue trajectory generation under the orchestration of the scene manager, given predefined roles and an initial scenario. The right part presents the evaluation of the actor models and the scene manager based on the generated trajectories.

### 3.1.2 Immersive Messaging Configuration

AdaMARP adopts a structured message representation that extends prior role-playing formats beyond plain dialogue [Shao et al., 2023, Chen et al., 2023, Li et al., 2023, Wang et al., 2024, Zhou et al., 2024, Lu et al., 2024, Zhou et al., 2024] and action-augmented interaction [Tu et al., 2024, He et al., 2025]. Following recent practice of explicitly modeling internal states (e.g., reasoning or thoughts) [Wang et al., 2025, Yang et al., 2025], we further introduce **environment-aware** descriptions as a first-class component of each turn.

Concretely, environmental signals play two complementary roles in our setting. First, they provide *atmospheric grounding*, using environmental cues to evoke or amplify character emotions. Second, they enable *interaction dynamics* between characters and their surroundings, where character actions update the environment state and external events (e.g., a sound behind a door) influence subsequent decisions and dialogue. To operationalize this design, we adopt a unified message format that interleaves Thought, Action, Environment, and Speech: internal thoughts are enclosed in square brackets [], physical actions in parentheses (), environment states or changes in angle brackets <>, and unmarked text denotes Speech.

### 3.1.3 Adaptive Role-Playing Framework

We propose the **AdaMARP framework** to support role-playing episodes in which the active speaker, the scene, and even the cast may change over time. The framework models three interacting agents. The **Actor Model**  $\mathcal{A}$  generates messages for all non-user characters, including the main character and any auxiliary roles. The **User Model**  $\mathcal{U}$  represents the user side of the interaction and is simulated by an LLM by default, while remaining replaceable by a human user. The **Scene Manager**  $\mathcal{S}$  performs high-level orchestration by tracking dialogue history and role/scene states.

Each role (including the user) is associated with a structured profile and a *scene-specific motivation*, which conditions behavior in the current context. At each step, the Scene Manager selects an action from a discrete control space  $\mathcal{M} = \{\text{init\_scene}, \text{pick\_speaker}, \text{switch\_scene}, \text{add\_role}, \text{end}\}$ . The first action issued in an episode is always `init_scene`, which provides the initial scene description and bootstraps the interaction history. For every action

---

**Algorithm 1:** Adaptive Role-Playing Framework

---

**Input:** Initial role set  $\mathcal{R}_0$ , role profiles  $\{\mathcal{P}_i\}$ , role motivations  $\{\Omega_i\}$   
**Output:** Role-Playing interaction trajectory  $\mathcal{H}$   
**Agents:** Scene Manager  $\mathcal{S}$ , Actor Model  $\mathcal{A}$ , User Model  $\mathcal{U}$ ;  
Initialize dialogue history  $\mathcal{H}_0 \leftarrow \emptyset$ ;  
 $t \leftarrow 0$ ;  
 $\mathcal{S}$  outputs  $m_0 = \text{init\_scene}$  with initial scene description;  
Append action output to  $\mathcal{H}_1$ ;  
 $t \leftarrow 1$ ;  
**while** *True* **do**  
    Construct interaction state  $\mathcal{G}_t = (\mathcal{R}_t, \mathcal{H}_t)$ ;  
     $\mathcal{S}$  selects action  $m_t \in \mathcal{M}$  with rationale;  
    **if**  $m_t = \text{pick\_speaker}$  **then**  
        Select role  $r_k \in \mathcal{R}_t$ ;  
        Generate in-character response using  $\mathcal{A}$  or  $\mathcal{U}$  conditioned on  $\mathcal{P}_k, \Omega_k, \{\mathcal{P}_j\}_{j \neq k}$ , and  $\mathcal{H}_t$ ;  
        Append response to  $\mathcal{H}_{t+1}$ ;  
    **else if**  $m_t = \text{switch\_scene}$  **then**  
        Append new scene description to  $\mathcal{H}_{t+1}$ ;  
    **else if**  $m_t = \text{add\_role}$  **then**  
        Instantiate new role  $r_{n_t+1}$  with profile  $\mathcal{P}_{n_t+1}$  and motivation  $\Omega_{n_t+1}$ ;  
         $\mathcal{R}_{t+1} \leftarrow \mathcal{R}_t \cup \{r_{n_t+1}\}$ ;  
        Append action output to  $\mathcal{H}_{t+1}$ ;  
    **else if**  $m_t = \text{end}$  **then**  
        **break**;  
     $t \leftarrow t + 1$ ;  
**return**  $\mathcal{H}_t$ ;

---

$m \in \mathcal{M}$ ,  $\mathcal{S}$  outputs an explicit rationale. When  $m = \text{switch\_scene}$ ,  $\mathcal{S}$  additionally produces an updated scene description; when  $m = \text{add\_role}$ , it specifies the new role's name, profile, and motivation; when  $m = \text{pick\_speaker}$ , it designates which role (including the user) should produce the next in-character message. The complete pseudo-code of the AdaMARP framework is presented in Algorithm 1, while additional implementation details and prompt templates for  $\mathcal{A}$ ,  $\mathcal{U}$ , and  $\mathcal{S}$  are provided in Appendix C.

## 3.2 Dataset Curation

Grounded in the above design principles, we curate two datasets to support the AdaMARP framework: **AdaRPSet** for training the Actor Model (Section 3.2.1) and **AdaSMSet** for supervising the Scene Manager (Section 3.2.2). Section 3.4 further presents an evaluation framework tailored to assess both components within AdaMARP.

### 3.2.1 Dataset for Actor Model

To satisfy the immersive messaging principle in Section 3.1.2, we construct **AdaRPSet**, a two-part corpus: **AdaRPSet-Extracted**, grounded in literary texts to teach the unified message format and more human-like behaviors, and **AdaRPSet-Synthesis**, designed to cover dynamic demands that are scarce in extracted data, such as scene transitions and on-the-fly character introduction.

**AdaRPSet-Extracted.** Following CoSER [Wang et al., 2025], we select 81 representative books from Goodreads' Best Books Ever list (Table 19) and obtain their full texts<sup>2</sup>. We build AdaRPSet-Extracted via three stages: **Chunking**, **LLM-based Extraction**, and **LLM-based Profile Generation**. We segment each book by chapters when available and merge

<sup>2</sup>[https://www.goodreads.com/list/show/1.Best\\_Books\\_Ever](https://www.goodreads.com/list/show/1.Best_Books_Ever)

adjacent segments into chunks under a size budget. We then prompt a strong LLM to identify coherent scenes and extract multi-character interaction trajectories (without restricting to two speakers [Yu et al., 2025]), emitting each trajectory *directly* in our unified messaging format (including *thought*, *action*, and *environment* when applicable), grounded in the source text. Finally, we aggregate scenes per character and prompt the LLM to synthesize a seven-dimensional profile (Section 3.1.1). Prompts and pipeline details are provided in Appendix D.1.

**AdaRPSet-Synthesis.** While AdaRPSet-Extracted supports well-formed messages conditioned on profiles and scenes, it is less robust to *dynamic* role-playing (e.g., scene switching and character insertion) and offers limited interaction diversity. We therefore construct AdaRPSet-Synthesis by prompting an advanced LLM to generate plot-level trajectories. Each trajectory specifies an initial scenario, a main character plus additional characters (each with a profile and initial motivation), and a multi-turn interaction in our unified format. It also includes *scene manager* control messages (e.g., Action=add\_role, Action=switch\_scene) to explicitly model role addition and scene transitions; we exclude pick Speaker to avoid turn-selection artifacts. We enforce at least one scene switch and one role addition per trajectory. We instantiate 20 themes with 50 trajectories each, using 45/5 for train/test (AdaptiveBench). The complete synthesis prompts and configurations are detailed in Appendix D.2, after which the resulting trajectories are standardized into a unified training-sample format as described in Appendix D.3.

### 3.2.2 Dataset for Scene Manager

To explicitly supervise high-level control in AdaMARP, we construct a dedicated dataset for the Scene Manager, **AdaSMS**. We build it on top of **AdaRPSet-Synthesis**, whose trajectories already include scene-manager actions for add\_role, switch\_scene, and end. We then add the remaining key supervision for  $\mathcal{S}$ —**speaker selection**—by inserting pick Speaker messages between consecutive character turns and prompting a strong instruction-following language model to generate the corresponding Reason field. Construction details and prompts are provided in Appendix D.4.

- ◆ We provide detailed statistical analyses of both AdaRPSet and AdaSMS in Appendix E.

## 3.3 Actor Model and Scene Manager Training

We train AdaMARP using supervised fine-tuning on two specialized corpora. The Actor Model  $\mathcal{A}$  is fine-tuned on **AdaRPSet** (comprising both Extracted and Synthesis subsets), where each training sample follows the unified role-playing message format detailed in Table 31. The Scene Manager  $\mathcal{S}$  is trained on **AdaSMS**, which provides explicit supervision for control decisions. Specifically,  $\mathcal{S}$  learns to predict the next action from the set  $\mathcal{M}$  (e.g., pick Speaker, switch\_scene, add\_role, end) along with its requisite arguments—such as reason, new\_scene, or new\_role\_\*—conditioned on the interaction history (See Table 32).

## 3.4 AdaptiveBench: Evaluation Framework

### 3.4.1 Dialogue Simulation

We introduce **AdaptiveBench**, a simulation framework designed to evaluate both the Actor Model and the Scene Manager by generating full interaction trajectories. We derive 100 evaluation seeds from the held-out split of AdaRPSet-Synthesis (5 instances per topic across 20 topics), preserving the initial role configurations and scene descriptions. For each seed, we simulate an episode involving three agents: an LLM-based user simulator  $U$ , the Actor Model  $\mathcal{A}$  portraying all non-user roles, and the Scene Manager  $\mathcal{S}$  issuing control actions (e.g., pick Speaker, switch\_scene). At every step,  $\mathcal{S}$  orchestrates the flow, and the designated speaker (either  $U$  or  $\mathcal{A}$ ) generates the response. Each simulation runs for a fixed horizon of  $T = 20$  dialogue turns (excluding  $\mathcal{S}$  messages), yielding 100 complete trajectories for subsequent evaluation.

### 3.4.2 Trajectory Evaluation

Since AdaMARP introduces a structured message format (incorporating environmental information) and dynamic states (e.g., scene transitions, role additions), existing role-playing evaluation frameworks [Brahman et al., 2021, Zhou et al., 2024, Yuan et al., 2024, Tu et al., 2024, Ahn et al., 2024, Samuel et al., 2025, Zhang et al., 2025, Lu et al., 2025] are

insufficient to capture these dimensions. We therefore adopt a *trajectory-level* protocol tailored for adaptive role-playing in AdaptiveBench, evaluating both the Actor Model and the Scene Manager on complete simulated dialogues.

**Actor Model evaluation.** For each simulated trajectory  $\tau = \{m_t\}_{t=1}^T$ , we provide an LLM-as-Judge with the role set  $\mathcal{R}$  (profiles), motivations  $\Omega$ , the initial scene  $\mathcal{E}_0$ , and the full dialogue trajectory  $\tau$ . The judge assigns independent 0–10 scores to each sub-metric under five dimensions: (I) **Character Consistency**: Internal Coherence, Speaking Style Fidelity, Language Fluency & Human-likeness, Identity & Profile Fidelity, Motivation & Value Stability; (II) **Environmental Grounding**: Env Awareness, Env Utilization; (III) **Interpersonal Interaction**: Contextual Responsiveness, Relationship Awareness; (IV) **Narrative Progression**: Attractiveness, Stability; (V) **Instruction Compliance**: Compliance. We report all sub-metric scores directly; the complete rubric and judge prompts are provided in Appendix F.1.

**Scene Manager evaluation.** We evaluate the Scene Manager with the same inputs as Actor evaluation (i.e.,  $\mathcal{R}$ ,  $\mathcal{M}$ ,  $\mathcal{E}_0$ , and the full trajectory  $\tau$ ), and ask an LLM to output four independent 0–10 scores: (I) **Scene Understanding**; (II) **Speaker Discipline**; (III) **Role Introduction Judgment**; (IV) **Overall Assessment**. The complete scoring rubric and judge prompts are provided in Appendix F.2.

- ◆ To complement the above dataset construction and evaluation framework, we provide a detailed analysis of token consumption across both data curation and AdaptiveBench evaluation in Appendix N.

## 4 Experiment

### 4.1 Experimental Settings

**Models.** We evaluate a diverse set of proprietary and open-source large language models. Our proprietary models include GPT-4o-mini [OpenAI et al., 2024], GPT-5-Chat [OpenAI, 2025], Gemini-2.5-Pro [Comanici et al., 2025], Claude Sonnet 4.5 [Anthropic, 2024], and Doubao-1.5-Pro-Character [ByteDance, 2024]. For open-source baselines, we consider three categories. (i) *Roleplay-oriented models*: Index-1.9B-Character<sup>3</sup>. (ii) *Instruction-tuned models*: Qwen2.5-7B/14B/72B-Instruct [Qwen Team et al., 2025] and Llama-3.1-70B-Instruct [Grattafiori et al., 2024]. (iii) *Reasoning-focused models*: Qwen3-14B and QwQ-32B [Yang et al., 2025]. To explicitly examine whether role-playing fine-tuning can effectively steer *base* models, we additionally include Llama-3.1-8/70B Grattafiori et al. [2024] as controlled baselines.

**Datasets and Training.** We fine-tune the Actor Model on AdaRPSet and the Scene Manager on AdaSMSSet, both curated in Section 3.2. All models are trained with full fine-tuning; detailed configurations and hyperparameters are provided in Appendix G.

**Baselines.** We focus on *general-purpose* role-playing and thus exclude character-specific methods (e.g., Neeko and HyCoRA) and single-turn settings (e.g., RoleLLM). We compare with three recent open-source baselines: BeyondDialogue (Beyond), Crab, and CoSER, which model increasingly rich signals (speech → action → thought); CoSER is the closest due to its multi-role support. We design two sets of inference-time system prompts—Basic and Enhance—for both the Actor Model and the Scene Manager (See Appendix H). Unless otherwise specified, we report results using the Enhance prompt for both components.

**Evaluation Metrics.** For the Actor Model, we primarily use AdaptiveBench to evaluate all models (open-source, proprietary, and baselines) on the five dimensions and twelve sub-metrics defined in Section 3.4.2. We further report results on two public role-playing evaluation methods, CharacterArena [Ye et al., 2025] and CharacterBench [Zhou et al., 2024], to examine performance under different prompts and metric designs. For the Scene Manager, since it is newly introduced in this work, we evaluate it only with AdaptiveBench. Details of CharacterArena and CharacterBench are provided in Appendix I. For AdaptiveBench, we report the mean score and standard deviation for each metric, where the standard deviation is computed over all evaluation samples within the benchmark.

<sup>3</sup><https://huggingface.co/IndexTeam/Index-1.9B-Character>

	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	IC	Average
Models	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	IC	
<b>Close-source LLMs</b>													
GPT-4o-mini	8.99±0.10	8.19±0.39	8.89±0.31	9.50±0.50	8.88±0.57	8.71±0.48	7.83±0.45	9.20±0.40	8.15±0.36	8.17±0.43	9.24±0.43	9.64±0.56	8.78
GPT-5-Chat	9.09±0.99	8.84±0.92	8.94±0.95	9.85±1.01	9.03±1.04	9.03±1.03	8.39±1.05	9.43±1.07	8.52±1.00	8.90±0.97	9.47±1.08	<b>9.70±0.46</b>	9.10
Gemini-2.5-Pro	7.97±0.41	7.57±0.60	7.57±0.55	8.92±0.46	8.02±0.53	8.04±0.62	7.17±0.68	8.30±0.57	7.87±0.72	8.27±0.58	7.88±0.70	9.22±0.50	8.07
Claude Sonnet 4.5	<b>9.35±0.54</b>	<b>8.97±0.22</b>	<b>8.99±0.41</b>	<b>9.94±0.28</b>	<b>9.27±0.66</b>	<b>9.13±0.58</b>	<b>8.39±0.63</b>	<b>9.78±0.44</b>	<b>8.90±0.48</b>	<b>9.23±0.53</b>	<b>9.65±0.65</b>	<b>9.61±0.49</b>	<b>9.27</b>
Doubao-1.5-Pro-Character	8.65±0.48	7.82±0.48	8.74±0.46	9.17±0.53	8.42±0.71	8.21±0.65	7.26±0.67	8.91±0.40	7.86±0.53	7.91±0.49	8.86±0.55	9.28±0.71	8.42
<b>Open-source LLMs</b>													
<b>&lt;14B</b>													
Index-1.9B-Character	5.62±1.46	4.63±1.35	6.81±1.08	6.15±1.59	6.49±1.39	6.17±1.28	4.82±1.32	6.84±1.35	5.43±1.36	5.13±1.27	6.69±1.41	5.47±2.61	5.85
Qwen2-7B-Beyond-RP*	5.32±1.99	4.77±2.13	6.53±1.81	6.05±2.38	6.13±2.05	5.75±1.66	4.54±1.68	6.39±2.09	5.34±1.92	5.07±1.94	6.29±1.96	5.12±2.81	5.61
Qwen2.5-7B-Instruct	8.61±0.99	7.70±0.91	8.70±0.96	9.03±1.02	8.35±1.06	8.24±1.06	7.33±0.95	8.81±0.97	7.80±0.89	7.73±0.96	8.89±0.97	9.23±0.72	8.37
Qwen2.5-7B-Instruct-Beyond	6.38±1.75	6.02±1.71	7.42±1.21	7.52±1.73	7.10±1.45	6.07±1.39	4.78±1.34	7.48±1.41	6.56±1.44	5.91±1.45	7.37±1.47	6.71±2.95	6.61
Qwen2.5-7B-Instruct-Crab	6.68±1.28	5.70±1.29	7.50±0.94	7.45±1.42	7.34±1.18	6.95±1.13	5.66±1.16	7.45±1.37	6.39±1.18	6.04±1.45	7.62±1.08	7.72±1.79	6.88
Qwen2.5-7B-Instruct-CoSER	8.37±0.66	7.52±0.69	8.51±0.58	8.97±0.67	8.37±0.76	7.57±0.79	6.51±0.83	8.69±0.54	7.78±0.52	7.51±0.72	8.69±0.63	9.33±0.55	8.15
<b>Qwen2.5-7B-Instruct-Ours</b>	<b>8.86±0.92</b>	<b>8.28±0.97</b>	<b>8.83±0.94</b>	<b>9.51±1.08</b>	<b>8.81±1.07</b>	<b>8.54±1.01</b>	<b>7.65±0.90</b>	<b>8.99±0.94</b>	<b>8.11±0.94</b>	<b>8.38±0.98</b>	<b>9.04±0.99</b>	<b>9.70±0.46</b>	<b>8.72</b>
Llama-3.1-8B-Crab	7.69±0.90	6.91±0.90	8.07±0.72	8.48±0.94	8.24±0.92	7.43±0.80	6.34±0.85	8.50±0.84	7.42±0.84	7.11±0.87	8.45±0.95	7.54±1.98	7.68
Llama-3.1-8B-CoSER*	8.66±0.61	7.71±0.61	8.72±0.45	9.06±0.61	8.68±0.79	7.72±0.78	6.71±0.85	9.00±0.56	7.98±0.59	7.51±0.76	8.97±0.51	9.16±1.06	8.32
<b>Llama-3.1-8B-Ours</b>	<b>9.00±0.00</b>	<b>8.39±0.49</b>	<b>8.93±0.26</b>	<b>9.69±0.48</b>	<b>9.08±0.63</b>	<b>8.77±0.49</b>	<b>7.84±0.44</b>	<b>9.26±0.46</b>	<b>8.36±0.52</b>	<b>8.52±0.50</b>	<b>9.39±0.49</b>	<b>9.49±0.54</b>	<b>8.89</b>
<b>14-72B</b>													
Qwen2.5-14B-Instruct	8.90±0.71	7.99±0.87	8.90±0.33	9.32±1.06	8.74±1.07	8.75±0.52	7.78±0.52	9.03±0.36	8.04±0.60	8.17±0.49	9.17±0.58	9.57±0.90	8.70
<b>Qwen2.5-14B-Instruct-Ours</b>	<b>8.98±0.14</b>	<b>8.34±0.47</b>	<b>8.96±0.20</b>	<b>9.66±0.47</b>	<b>8.94±0.56</b>	<b>8.69±0.46</b>	<b>7.77±0.42</b>	<b>9.10±0.30</b>	<b>8.18±0.43</b>	<b>8.61±0.49</b>	<b>9.11±0.37</b>	<b>9.70±0.46</b>	<b>8.84</b>
Qwen2.5-72B-Instruct	8.93±0.26	8.10±0.33	8.96±0.20	9.40±0.53	8.72±0.66	8.61±0.55	7.66±0.51	9.13±0.39	8.14±0.46	8.05±0.46	9.17±0.45	<b>9.71±0.45</b>	8.71
<b>Qwen2.5-72B-Instruct-Ours</b>	<b>8.95±0.22</b>	<b>8.42±0.49</b>	<b>8.86±0.34</b>	<b>9.67±0.47</b>	<b>8.79±0.56</b>	<b>8.55±0.52</b>	<b>7.71±0.45</b>	<b>9.14±0.37</b>	<b>8.26±0.51</b>	<b>8.62±0.48</b>	<b>9.09±0.41</b>	<b>9.64±0.48</b>	8.81
Llama-3.1-70B-Instruct	8.91±0.90	8.08±0.89	8.88±0.91	9.36±1.06	8.79±1.08	8.39±1.03	7.58±1.01	9.07±0.98	8.14±0.93	8.39±0.98	9.10±0.99	9.49±1.09	8.68
Llama-3.1-70B-CoSER*	8.90±0.30	7.98±0.35	8.88±0.32	9.41±0.53	8.77±0.65	8.14±0.65	7.19±0.67	9.09±0.32	8.14±0.40	8.07±0.43	9.08±0.37	9.65±0.48	8.61
<b>Llama-3.1-70B-Ours</b>	<b>8.90±0.71</b>	<b>8.14±0.82</b>	<b>8.83±0.75</b>	<b>9.42±0.84</b>	<b>8.99±1.11</b>	<b>8.55±1.00</b>	<b>7.63±0.93</b>	<b>9.05±0.89</b>	<b>8.24±0.86</b>	<b>8.37±0.99</b>	<b>9.21±0.78</b>	<b>9.05±0.90</b>	8.70
Qwen3-14B	8.91±1.32	8.64±1.31	8.67±1.31	9.67±1.45	8.91±1.42	8.85±1.36	8.19±1.33	9.05±1.42	8.22±1.31	8.73±1.34	9.25±1.45	9.43±1.10	8.88
QwQ-32B	<b>9.19±0.39</b>	<b>8.99±0.22</b>	<b>8.89±0.34</b>	<b>9.96±0.20</b>	<b>9.11±0.51</b>	<b>9.15±0.50</b>	<b>8.72±0.69</b>	<b>9.35±0.54</b>	<b>8.43±0.51</b>	<b>9.17±0.43</b>	<b>9.45±0.52</b>	<b>9.26±0.64</b>	<b>9.14</b>

**Table 2.** Actor model evaluation results on **AdaptiveBench**, with GPT-5-Chat as the judge model. **Bold** indicates the best performance within the same model scale, and underline indicates the second-best performance within the same scale. Models marked with \* are evaluated using their officially released checkpoints without additional training, while unmarked baselines are re-trained under our experimental setup. ± denotes the standard deviation computed over all samples.

## 4.2 Main Results

### 4.2.1 Actor Model

Following the experimental setup described above, we evaluate all Actor Models (including open-/closed-source LLMs and re-trained baselines) on AdaptiveBench with a unified configuration: the User Model is simulated by Doubao-1.5-Pro-Character and the Manager Model is simulated by GPT-4o-mini, both equipped with our Enhance system prompts. We then use GPT-5-Chat as the judge model for trajectory-level scoring (we ablate different judge models in Appendix K). The results are shown in Table 2. Since a score of 5 denotes a neutral/acceptable level (see Table 34~37), we draw the following observations.

- (1) **Strong LLMs, instruction-tuned models, and reasoning models excel in adaptive role-playing.** Claude Sonnet 4.5 and GPT-5-Chat lead the proprietary tier, while the open-source reasoning model QwQ-32B is competitive. Notably, Qwen2.5-7B-Instruct already surpasses Gemini-2.5-Pro overall, and Qwen2.5-14B-Instruct approaches GPT-4o-mini; scaling Qwen2.5 to 72B brings diminishing returns, suggesting a bottleneck for standard instruction-tuned models. Among 14B models, the reasoning-oriented Qwen3-14B achieves the best overall performance. We further analyze the impact of prompting strategies by comparing Basic and Enhance settings in Appendix M.1, where we find that the simpler Basic prompting mostly outperforms Enhance, a counterintuitive result suggesting that adaptive role-playing benefits more from flexibility than from overly restrictive guidance.
- (2) **Existing methods struggle in the adaptive setting and may harm instruction-tuned base models.** Adaptive role-playing requires integrating Thought–Action–Speech–Environment and handling dynamic scene/role shifts. Accordingly, specialized models (e.g., Index-1.9B-Character) perform poorly, barely above the neutral baseline. More strikingly, fine-tuning a strong Qwen2.5-7B-Instruct base (8.37) on Beyond, Crab, or CoSER *reduces* performance (to 6.61, 6.88, and 8.15). We attribute this to (i) *format mismatch* (missing thought/environment signals), (ii) *interaction mismatch* (dyadic chat vs. multi-role coordination), and (iii) *distribution mismatch* (few scene transitions or role reassessments), which together make models brittle under dynamic trajectories.

Models	Scene Understanding	Speaker Discipline	Role Introduction Judgment	Overall Assessment
<b>Close-source LLMs</b>				
GPT-4o-mini	<u>7.64±1.13</u>	<u>8.55±0.87</u>	7.18±1.26	7.64±1.01
GPT-5-Chat	<u>8.03±0.75</u>	8.15±1.30	7.78±1.02	<u>7.90±0.90</u>
Claude Sonnet 4.5	<u>8.21±0.57</u>	<u>8.62±0.83</u>	<u>8.05±0.85</u>	<u>8.17±0.57</u>
Doubaot-1.5-Pro-Character	7.67±0.68	7.39±1.35	<u>7.95±1.12</u>	7.51±0.91
<b>Open-source LLMs</b>				
Qwen2.5-7B-Instruct	7.55±0.89	8.35±1.08	7.25±1.17	7.52±0.87
<b>Qwen2.5-7B-Instruct-Ours</b>	<b>7.96±0.71</b>	<u>8.21±1.11</u>	8.00±1.04	7.93±0.71
Qwen2.5-14B-Instruct	7.65±0.86	8.42±1.12	7.34±1.16	7.63±0.84
<b>Qwen2.5-14B-Instruct-Ours</b>	<u>8.23±0.44</u>	8.25±0.84	<b>8.64±0.59</b>	<u>8.37±0.52</u>
Qwen3-14B	7.81±0.66	8.49±0.84	7.60±1.10	7.80±0.69
Qwen2.5-72B-Instruct	7.86±0.53	<u>8.51±0.76</u>	7.65±0.99	7.82±0.61
Llama-3.1-70B-Instruct	7.75±0.75	8.33±0.99	7.85±1.19	7.79±0.84
QwQ-32B	<b>8.48±0.54</b>	<b>8.92±0.70</b>	7.97±0.71	<b>8.41±0.55</b>

**Table 3.** Scene Manager evaluation results on **AdaptiveBench**, with GPT-5-Chat as the judge model. **Bold** indicates the best performance within the same model scale, and underline indicates the second-best performance within the same scale.  $\pm$  denotes the standard deviation computed over all samples.

**(3) AdaRPSet yields substantial and robust improvements across model scales and protocols.** Fine-tuning with AdaRPSet consistently strengthens the Actor Model under AdaptiveBench, yielding broad improvements across sub-metrics (CC/EG/II/NP/IC) and generalizing across model scales (7B–72B) and backbones (Qwen2.5 vs. Llama-3.1). Concretely, Qwen2.5-7B-Instruct-Ours improves the overall score from 8.37 to 8.72 (+4.2%), while Qwen2.5-14B-Instruct-Ours and Qwen2.5-72B-Instruct-Ours rise to 8.84 (+1.6%) and 8.81 (+1.1%), respectively. Notably, Llama-3.1-8B-Ours achieves 8.89, outperforming Qwen3-14B (8.88), GPT-4o-mini (8.78), and several 70B/72B baselines, demonstrating that our training recipe elicits strong adaptive capabilities even at smaller scales. Importantly, these improvements are also robust across evaluation protocols: beyond automatic judging, human evaluation results (Appendix I.3) confirm the same trend, and our method *mostly* outperforms baselines on CharacterArena and CharacterBench (Appendix J.1). We further provide qualitative case studies in Appendix O to illustrate these gains in practice.

#### 4.2.2 Scene Manager

We evaluate the **Scene Manager** on AdaptiveBench by fixing the User and Actor models (Doubaot-1.5-Pro-Character) and varying only the manager. We use the Enhance system prompt and GPT-5-Chat as the judge (ablations in Appendix L). Results in Table 3 (where 5 point denotes neutral performance) reveal the following:

**(1) Standard instruction tuning yields compressed performance, while reasoning models break the ceiling.** Most instruction-tuned models cluster in a narrow Overall Assessment range (7.5–7.9), suggesting that general-purpose tuning does not explicitly optimize scene coordination skills (e.g., state tracking and turn-taking). For instance, Qwen2.5-7B/14B/72B and Llama-3.1-70B improve only marginally and plateau around 7.8, forming a practical ceiling for this class. In contrast, stronger proprietary models (Claude Sonnet 4.5, GPT-5-Chat) perform consistently better, and the reasoning-oriented QwQ-32B reaches 8.41, indicating that reasoning-centric training better supports high-level orchestration than standard instruction tuning alone. Moreover, Appendix M.2 shows that Enhance prompts (with stronger constraints) improve Scene Manager performance, opposite to the trend for the Actor Model.

**(2) AdaSMSet substantially improves scene management and approaches the best open-source reasoning model.** Training on AdaSMSet yields clear gains for the Scene Manager. Qwen2.5-14B-Instruct-Ours reaches 8.37 overall, outperforming Claude Sonnet 4.5 (8.17) and improving over the base Qwen2.5-14B-Instruct (7.63) by +0.74 (+9.7%). *Speaker Discipline* drops slightly, which is expected since AdaSMSet is derived from AdaRPSet-Synthesis and partially inherits GPT-5-Chat’s turn-taking patterns (with pick\_speaker as the most frequent action). Importantly, *Scene Understanding* and *Role Introduction Judgment* improve substantially, yielding a strong net gain and bringing performance close to the best open-source model here (QwQ-32B, 8.41). A similar trend holds for Qwen2.5-7B-Instruct: the overall score rises from 7.52 to 7.93 (+0.41), with consistent gains on *Scene Understanding* and *Role Introduction Judgment*, while *Speaker Discipline* remains comparable.

	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	IC	IC	Average
<b>Models</b>														
<b>Owen2.5-7B-Instruct</b>														
Base	8.61±0.99	7.70±0.91	8.70±0.96	9.03±1.02	8.35±1.06	8.24±1.06	7.33±0.95	8.81±0.97	7.80±0.89	7.73±0.96	8.89±0.97	9.23±0.72	8.37	
+Extracted	8.67±0.47	8.00±0.42	8.82±0.38	9.29±0.45	8.40±0.61	8.24±0.56	7.36±0.64	8.89±0.38	7.96±0.51	7.96±0.51	8.87±0.45	9.42±0.49	8.49	
+Extracted+Synthesis	8.86±0.92	8.28±0.97	8.83±0.94	9.51±1.08	8.81±1.07	8.54±1.01	7.65±0.90	8.99±0.94	8.11±0.94	8.38±0.98	9.04±0.99	9.70±0.46	8.72	
<b>Llama-3.1-8B</b>														
+Extracted	7.97±1.24	7.49±0.98	7.50±1.57	8.68±1.14	8.30±1.12	7.75±1.13	6.83±1.08	8.36±1.14	7.55±1.08	7.48±1.13	8.07±1.25	6.45±1.75	7.70	
+Extracted+Synthesis	9.00±0.00	8.39±0.49	8.93±0.26	9.69±0.48	9.08±0.63	8.77±0.49	7.84±0.44	9.26±0.46	8.36±0.52	8.52±0.50	9.39±0.49	9.49±0.54	8.89	

**Table 4.** Actor model ablation results on **AdaptiveBench** (judge: GPT-5-Chat). “Extracted” denotes training with AdaRPSet-Extracted, and “Extracted+Synthesis” denotes training with full AdaRPSet (AdaRPSet-Extracted + AdaRPSet-Synthesis).  $\pm$  denotes the standard deviation computed over all samples.

### 4.3 Ablation Study

To quantify the contribution of each AdaRPSet component, we train Actor Models with (i) **AdaRPSet-Extracted** and (ii) the full AdaRPSet (*Extracted+Synthesis*). We evaluate our method on AdaptiveBench using GPT-5-Chat as the judge, and the results are shown in Table 4.

**AdaRPSet-Extracted alone provides consistent gains and improves stability.** On Qwen2.5-7B-Instruct, training with extracted data improves the overall score from 8.37 to 8.49 (+0.12). The gains are broad-based across most sub-metrics (e.g., CC and IC), indicating that AdaRPSet-Extracted effectively teaches the model to follow the unified Thought–Action–Speech–Env protocol and to better maintain character consistency during interaction. For the few dimensions where the mean does not increase (e.g., EA/RA/STB), we observe noticeably smaller variance, suggesting more stable behavior across different trajectories.

**AdaRPSet-Synthesis yields further improvements and is crucial for generalization.** When we incorporate the synthesis data, Qwen2.5-7B-Instruct improves to 8.72 overall, i.e., +0.23 over *Extracted*-only and +0.35 over the base model (8.37 → 8.49 → 8.72). Gains are more pronounced on interaction- and narrative-related dimensions (e.g., II/NP), consistent with AdaRPSet-Synthesis containing more adaptive trajectories with scene transitions and role re-assignment. This highlights the complementarity of the two subsets: extracted data strengthens format alignment and core role-playing skills, while synthesis data improves adaptivity under dynamic trajectories. This trend generalizes to other model families. For Llama-3.1-8B, we omit the base result because the vanilla model does not reliably follow our structured protocol. Training with *AdaRPSet-Extracted* alone is insufficient (7.70 overall; IC: 6.45), whereas adding *AdaRPSet-Synthesis* substantially boosts performance to 8.89 overall (+1.19) and IC to 9.49, demonstrating strong generalization beyond the Qwen family.

## 5 Conclusion

We present **AdaMARP**, a general role-playing framework that boosts immersion and adaptability with (i) an environment-aware format interleaving thought, action, environment, and speech, and (ii) a discrete-action Scene Manager for multi-character coordination, scene transitions, and dynamic role addition. We release AdaRPSet, AdaSMS, and AdaptiveBench for trajectory-level evaluation. Experiments show consistent improvements in role consistency, narrative coherence, and environmental grounding across model scales; our 8B Actor and 14B Scene Manager outperform several larger proprietary systems, pointing toward agentic, open-ended role-playing beyond fixed response patterns.

## Ethical Considerations

The extracted subset of AdaRPSet is derived from existing literary works. We emphasize that the dataset is used *exclusively for scientific research* and not for any commercial purpose. The data consists of transformed and abstracted representations rather than verbatim copyrighted text, and all rights to the original works remain with their respective copyright holders. We provide this dataset solely for academic use and disclaim responsibility for any misuse beyond its intended research scope.

## References

- [1] Zhen Tan, Jun Yan, I-Hung Hsu, Rujun Han, Zifeng Wang, Long T. Le, Yiwen Song, Yanfei Chen, Hamid Palangi, George Lee, Anand Iyer, Tianlong Chen, Huan Liu, Chen-Yu Lee, and Tomas Pfister. In prospect and retrospect: Reflective memory management for long-term personalized dialogue agents, 2025. URL <https://arxiv.org/abs/2503.08026>.
- [2] Jusheng Zhang, Yijia Fan, Kaitong Cai, Xiaofei Sun, and Keze Wang. Osc: Cognitive orchestration through dynamic knowledge alignment in multi-agent llm collaboration, 2025. URL <https://arxiv.org/abs/2509.04876>.
- [3] Yufan Zhuang, Xiaodong Yu, Jialian Wu, Ximeng Sun, Ze Wang, Jiang Liu, Yusheng Su, Jingbo Shang, Zicheng Liu, and Emad Barsoum. Self-taught agentic long context understanding, 2025. URL <https://arxiv.org/abs/2502.15920>.
- [4] Jensen Zhang, Ningyuan Liu, Yijia Fan, Zihao Huang, Qinglin Zeng, Kaitong Cai, Jian Wang, and Keze Wang. Llm-cas: Dynamic neuron perturbation for real-time hallucination correction, 2025. URL <https://arxiv.org/abs/2512.18623>.
- [5] Zhenhua Xu, Xixiang Zhao, Xubin Yue, Shengwei Tian, Changting Lin, and Meng Han. CTCC: A Robust and Stealthy Fingerprinting Framework for Large Language Models via Cross-Turn Contextual Correlation Backdoor. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 6978–7000, Suzhou, China, 2025. Association for Computational Linguistics. ISBN 979-8-89176-332-6. doi: 10.18653/v1/2025.emnlp-main.356.
- [6] Dezhong Kong, Shi Lin, Zhenhua Xu, Zhebo Wang, Minghao Li, Yufeng Li, Yilun Zhang, Hujin Peng, Xiang Chen, Zeyang Sha, Yuyuan Li, Changting Lin, Xun Wang, Xuan Liu, Ningyu Zhang, Chaochao Chen, Chunming Wu, Muhammad Khurram Khan, and Meng Han. A survey of llm-driven ai agent communication: Protocols, security risks, and defense countermeasures, 2025. URL <https://arxiv.org/abs/2506.19676>.
- [7] Zhenhua Xu, Meng Han, and Wenpeng Xing. EverTracer: Hunting Stolen Large Language Models via Stealthy and Robust Probabilistic Fingerprint. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 7019–7042, Suzhou, China, 2025. Association for Computational Linguistics. ISBN 979-8-89176-332-6. doi: 10.18653/v1/2025.emnlp-main.358.
- [8] Junxian Li, Beining Xu, Simin Chen, Jiatong Li, Jingdi Lei, Haodong Zhao, and Di Zhang. Iag: Input-aware backdoor attack on vlm-based visual grounding. *arXiv preprint arXiv:2508.09456*, 2025.
- [9] Zhenhua Xu, Xubin Yue, Zhebo Wang, Qichen Liu, Xixiang Zhao, Jingxuan Zhang, Wenjun Zeng, Wengpeng Xing, Dezhong Kong, Changting Lin, et al. Copyright protection for large language models: A survey of methods, challenges, and trends. *arXiv preprint arXiv:2508.11548*, 2025.
- [10] Murray Shanahan, Kyle McDonell, and Laria Reynolds. Role-Play with Large Language Models, 2023.
- [11] Yu-Min Tseng, Yu-Chao Huang, Teng-Yun Hsiao, Wei-Lin Chen, Chao-Wei Huang, Yu Meng, and Yun-Nung Chen. Two Tales of Persona in LLMs: A Survey of Role-Playing and Personalization. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 16612–16631. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.findings-emnlp.969.
- [12] Jiangjie Chen, Xiantao Wang, Rui Xu, Siyu Yuan, Yikai Zhang, Wei Shi, Jian Xie, Shuang Li, Ruihan Yang, Tinghui Zhu, Aili Chen, Nianqi Li, Lida Chen, Caiyu Hu, Siye Wu, Scott Ren, Ziquan Fu, and Yanghua Xiao. From Persona to Personalization: {A} Survey on Role-Playing Language Agents. 2024, 2024. doi: 10.48550/arXiv.2404.18231.
- [13] Xiantao Wang, Heng Wang, Yifei Zhang, Xinfeng Yuan, Rui Xu, Jen-tse Huang, Siyu Yuan, Haoran Guo, Jiangjie Chen, Shuchang Zhou, Wei Wang, and Yanghua Xiao. CoSER: Coordinating LLM-Based Persona Simulation of Established Roles, 2025.

- [14] Xiaoyan Yu, Tongxu Luo, Yifan Wei, Fangyu Lei, Yiming Huang, Hao Peng, and Liehuang Zhu. Neeko: Leveraging Dynamic LoRA for Efficient Multi-Character Role-Playing Agent. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 12540–12557. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.emnlp-main.697.
- [15] Shihao Yang, Zhicong Lu, Yong Yang, Bo Lv, Yang Shen, and Nayu Liu. HyCoRA: Hyper-Contrastive Role-Adaptive Learning for Role-Playing, 2025.
- [16] Yunfan Shao, Linyang Li, Junqi Dai, and Xipeng Qiu. Character-LLM: A Trainable Agent for Role-Playing. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13153–13187. Association for Computational Linguistics, 2023. URL <https://aclanthology.org/2023.emnlp-main.814/>.
- [17] Nuo Chen, Yan Wang, Haiyun Jiang, Deng Cai, Yuhan Li, Ziyang Chen, Longyue Wang, and Jia Li. Large Language Models Meet Harry Potter: A Dataset for Aligning Dialogue Agents with Characters. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 8506–8520. Association for Computational Linguistics, 2023. doi: 10.18653/v1/2023.findings-emnlp.570.
- [18] Cheng Li, Ziang Leng, Chenxi Yan, Junyi Shen, Hao Wang, Weishi Mi, Yaying Fei, Xiaoyang Feng, Song Yan, HaoSheng Wang, Linkang Zhan, Yaokai Jia, Pingyu Wu, and Haozhen Sun. ChatHaruhi: Reviving Anime Character in Reality via Large Language Model, 2023. URL <https://arxiv.org/abs/2308.09597v1>.
- [19] Noah Wang, Z.y. Peng, Haoran Que, Jiaheng Liu, Wangchunshu Zhou, Yuhan Wu, Hongcheng Guo, Ruitong Gan, Zehao Ni, Jian Yang, Man Zhang, Zhaoxiang Zhang, Wanli Ouyang, Ke Xu, Wenhao Huang, Jie Fu, and Junran Peng. RoleLLM: Benchmarking, Eliciting, and Enhancing Role-Playing Abilities of Large Language Models. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 14743–14777. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.findings-acl.878.
- [20] Jinfeng Zhou, Zhuang Chen, Dazhen Wan, Bosi Wen, Yi Song, Jifan Yu, Yongkang Huang, Pei Ke, Guanqun Bi, Libiao Peng, JiaMing Yang, Xiyao Xiao, Sahand Sabour, Xiaohan Zhang, Wenjing Hou, Yijia Zhang, Yuxiao Dong, Hongning Wang, Jie Tang, and Minlie Huang. CharacterGLM: Customizing social characters with large language models. In Franck Dernoncourt, Daniel Preoțiuc-Pietro, and Anastasia Shimorina, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 1457–1476, Miami, Florida, US, November 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.emnlp-industry.107. URL <https://aclanthology.org/2024.emnlp-industry.107/>.
- [21] Keming Lu, Bowen Yu, Chang Zhou, and Jingren Zhou. 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)*, pages 7828–7840. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.423.
- [22] Jinfeng Zhou, Yongkang Huang, Bosi Wen, Guanqun Bi, Yuxuan Chen, Pei Ke, Zhuang Chen, Xiyao Xiao, Libiao Peng, Kuntian Tang, Rongsheng Zhang, Le Zhang, Tangjie Lv, Zhipeng Hu, Hongning Wang, and Minlie Huang. CharacterBench: Benchmarking Character Customization of Large Language Models, 2024.
- [23] Quan Tu, Shilong Fan, Zihang Tian, Tianhao Shen, Shuo Shang, Xin Gao, and Rui Yan. CharacterEval: A Chinese Benchmark for Role-Playing Conversational Agent Evaluation. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11836–11850. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.acl-long.638.
- [24] Kai He, Yucheng Huang, Wenqing Wang, Delong Ran, Dongming Sheng, Junxuan Huang, Qika Lin, Jiaxing Xu, Wenqiang Liu, and Mengling Feng. Crab: A Novel Configurable Role-Playing LLM with Assessing Benchmark. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15030–15052. Association for Computational Linguistics, 2025. ISBN 979-8-89176-251-0. doi: 10.18653/v1/2025.acl-long.731.

- [25] Yiting Ran, Xintao Wang, Rui Xu, Xinfeng Yuan, Jiaqing Liang, Yanghua Xiao, and Deqing Yang. Capturing Minds, Not Just Words: Enhancing Role-Playing Language Models with Personality-Indicative Data. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Findings of the Association for Computational Linguistics: EMNLP 2024*, pages 14566–14576. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.findings-emnlp.853.
- [26] Bohao Yang, Dong Liu, Chenghao Xiao, Kun Zhao, Chen Tang, Chao Li, Lin Yuan, Yang Guang, and Chenghua Lin. Crafting Customisable Characters with LLMs: A Persona-Driven Role-Playing Agent Framework. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 20216–20240. Association for Computational Linguistics, 2025. ISBN 979-8-89176-335-7. doi: 10.18653/v1/2025.findings-emnlp.1100.
- [27] Yeyong Yu, Runsheng Yu, Haojie Wei, Zhanqiu Zhang, and Quan Qian. Beyond Dialogue: A Profile-Dialogue Alignment Framework Towards General Role-Playing Language Model. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11992–12022. Association for Computational Linguistics, 2025. ISBN 979-8-89176-251-0. doi: 10.18653/v1/2025.acl-long.586.
- [28] Zhenpeng Gao, Xiaofen Xing, and Xiangmin Xu. TailorRPA: A Retrieval-Based Framework for Eliciting Personalized and Coherent Role-Playing Agents in General Domain. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 5381–5412. Association for Computational Linguistics, 2025. ISBN 979-8-89176-335-7. doi: 10.18653/v1/2025.findings-emnlp.288.
- [29] Ke Ji, Yixin Lian, Linxu Li, Jingsheng Gao, Weiyuan Li, and Bin Dai. Enhancing Persona Consistency for LLMs’ Role-Playing using Persona-Aware Contrastive Learning. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Findings of the Association for Computational Linguistics: ACL 2025*, pages 26221–26238. Association for Computational Linguistics, 2025. ISBN 979-8-89176-256-5. doi: 10.18653/v1/2025.findings-acl.1344.
- [30] Haiming Qin, Jiwei Zhang, Wei Zhang, KeZhong Lu, Mingyang Zhou, Hao Liao, and Rui Mao. R-CHAR: A Metacognition-Driven Framework for Role-Playing in Large Language Models. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 26984–27002. Association for Computational Linguistics, 2025. ISBN 979-8-89176-332-6. doi: 10.18653/v1/2025.emnlp-main.1372.
- [31] Jing Ye, Rui Wang, Yuchuan Wu, Victor Ma, Feiteng Fang, Fei Huang, and Yongbin Li. CPO: Addressing Reward Ambiguity in Role-playing Dialogue via Comparative Policy Optimization. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 297–323. Association for Computational Linguistics, 2025. ISBN 979-8-89176-335-7. doi: 10.18653/v1/2025.findings-emnlp.18.
- [32] Cheng Liu, Yifei Lu, Fanghua Ye, Jian Li, Xingyu Chen, Feiliang Ren, Zhaopeng Tu, and Xiaolong Li. Cog-Dual: Enhancing Dual Cognition of LLMs via Reinforcement Learning with Implicit Rule-Based Rewards. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing*, pages 27295–27324. Association for Computational Linguistics, 2025. ISBN 979-8-89176-332-6. doi: 10.18653/v1/2025.emnlp-main.1389.
- [33] Faeze Brahman, Meng Huang, Oyvind Tafjord, Chao Zhao, Mrinmaya Sachan, and Snigdha Chaturvedi. “Let Your Characters Tell Their Story”: A Dataset for Character-Centric Narrative Understanding. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 1734–1752. Association for Computational Linguistics, 2021. doi: 10.18653/v1/2021.findings-emnlp.150.
- [34] Jinfeng Zhou, Zhuang Chen, Dazhen Wan, Bosi Wen, Yi Song, Jifan Yu, Yongkang Huang, Pei Ke, Guanqun Bi, Libiao Peng, JiaMing Yang, Xiyao Xiao, Sahand Sabour, Xiaohan Zhang, Wenjing Hou, Yijia Zhang, Yuxiao Dong, Hongning Wang, Jie Tang, and Minlie Huang. CharacterGLM: Customizing Social Characters with Large

- Language Models. In Franck Dernoncourt, Daniel Preoțiuc-Pietro, and Anastasia Shimorina, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*, pages 1457–1476. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.emnlp-industry.107.
- [35] Xinfeng Yuan, Siyu Yuan, Yuhang Cui, Tianhe Lin, Xintao Wang, Rui Xu, Jiangjie Chen, and Deqing Yang. Evaluating Character Understanding of Large Language Models via Character Profiling from Fictional Works. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 8015–8036. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.emnlp-main.456.
  - [36] Jaewoo Ahn, Taehyun Lee, Junyoung Lim, Jin-Hwa Kim, Sangdoo Yun, Hwaran Lee, and Gunhee Kim. TimeChara: Evaluating Point-in-Time Character Hallucination of Role-Playing Large Language Models. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 3291–3325. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.findings-acl.197.
  - [37] Vinay Samuel, Henry Peng Zou, Yue Zhou, Shreyas Chaudhari, Ashwin Kalyan, Tanmay Rajpurohit, Ameet Deshpande, Karthik R Narasimhan, and Vishvak Murahari. PersonaGym: Evaluating Persona Agents and LLMs. In Christos Christodoulopoulos, Tanmoy Chakraborty, Carolyn Rose, and Violet Peng, editors, *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 6999–7022. Association for Computational Linguistics, 2025. ISBN 979-8-89176-335-7. doi: 10.18653/v1/2025.findings-emnlp.368.
  - [38] Pinyi Zhang, Siyu An, Lingfeng Qiao, Yifei Yu, Jingyang Chen, Jie Wang, Di Yin, Xing Sun, and Kai Zhang. RolePlot: A Systematic Framework for Evaluating and Enhancing the Plot-Progression Capabilities of Role-Playing Agents. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 12337–12354. Association for Computational Linguistics, 2025. ISBN 979-8-89176-251-0. doi: 10.18653/v1/2025.acl-long.603.
  - [39] Junru Lu, Jiazheng Li, Guodong Shen, Lin Gui, Siyu An, Yulan He, Di Yin, and Xing Sun. RoleMRC: A Fine-Grained Composite Benchmark for Role-Playing and Instruction-Following. In Wanxiang Che, Joyce Nabende, Ekaterina Shutova, and Mohammad Taher Pilehvar, editors, *Findings of the Association for Computational Linguistics: ACL 2025*, pages 21008–21030. Association for Computational Linguistics, 2025. ISBN 979-8-89176-256-5. doi: 10.18653/v1/2025.findings-acl.1082.
  - [40] OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, and 262 others. Gpt-4 technical report, 2024. URL <https://arxiv.org/abs/2303.08774>.
  - [41] OpenAI. Gpt-5 system card. <https://cdn.openai.com/gpt-5-system-card.pdf>, Aug 2025. Version dated August 13, 2025.
  - [42] Gheorghe Comanici et al. Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities, 2025. URL <https://arxiv.org/abs/2507.06261>.
  - [43] Anthropic. Claude 4 model family. <https://www.anthropic.com>, 2024. Claude Sonnet 4.5.
  - [44] ByteDance. Doubao large language models. <https://www.volcengine.com>, 2024. Doubao-1.5-Pro-Character.
  - [45] Qwen Team et al. Qwen2.5 technical report, 2025. URL <https://arxiv.org/abs/2412.15115>.
  - [46] Aaron Grattafiori et al. The llama 3 herd of models, 2024. URL <https://arxiv.org/abs/2407.21783>.
  - [47] An Yang et al. Qwen3 technical report, 2025. URL <https://arxiv.org/abs/2505.09388>.
  - [48] Nathan Lambert, Jacob Morrison, Valentina Pyatkin, Shengyi Huang, Hamish Ivison, Faeze Brahman, Lester James V. Miranda, Alisa Liu, Nouha Dziri, Shane Lyu, Yuling Gu, Saumya Malik, Victoria Graf, Jena D. Hwang, Jiangjiang Yang, Ronan Le Bras, Oyvind Tafjord, Chris Wilhelm, Luca Soldaini, Noah A. Smith, Yizhong Wang,

- Pradeep Dasigi, and Hannaneh Hajishirzi. Tulu 3: Pushing frontiers in open language model post-training, 2025. URL <https://arxiv.org/abs/2411.15124>.
- [49] Dongsheng Chen, Zhiqi Huang, Xian Wu, Shen Ge, and Yuexian Zou. Towards joint intent detection and slot filling via higher-order attention. In *IJCAI*, pages 4072–4078, 2022.
- [50] Zhihong Zhu, Xuxin Cheng, Hao An, Zhichang Wang, Dongsheng Chen, and Zhiqi Huang. Zero-shot spoken language understanding via large language models: A preliminary study. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 17877–17883, 2024.
- [51] Zhihong Zhu, Xuxin Cheng, Zhiqi Huang, Dongsheng Chen, and Yuexian Zou. Towards unified spoken language understanding decoding via label-aware compact linguistics representations. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 12523–12531, 2023.
- [52] Zhihong Zhu, Xuxin Cheng, Zhiqi Huang, Dongsheng Chen, and Yuexian Zou. Enhancing code-switching for cross-lingual slu: A unified view of semantic and grammatical coherence. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 7849–7856, 2023.
- [53] Dongsheng Chen, Chaofan Tao, Lu Hou, Lifeng Shang, Xin Jiang, and Qun Liu. Litevl: Efficient video-language learning with enhanced spatial-temporal modeling. *arXiv preprint arXiv:2210.11929*, 2022.
- [54] Zhihong Zhu, Xuxin Cheng, Dongsheng Chen, Zhiqi Huang, Hongxiang Li, and Yuexian Zou. Mix before align: Towards zero-shot cross-lingual sentiment analysis via soft-mix and multi-view learning. In *Proc. of INTERSPEECH*, 2023.
- [55] Zhiqi Huang, Dongsheng Chen, Zhihong Zhu, and Xuxin Cheng. Mclf: A multi-grained contrastive learning framework for asr-robust spoken language understanding. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 7936–7949, 2023.
- [56] Hang Ding, Qiming Feng, Dongqi Liu, Qi Zhao, Tao Yao, Shuo Wang, Dongsheng Chen, Jian Li, Zhenye Gan, Jiangning Zhang, et al. Rolermbench & rolerm: Towards reward modeling for profile-based role play in dialogue systems. *arXiv preprint arXiv:2512.10575*, 2025.

## A Additional Clarifications on Dataset Comparison

This appendix provides detailed clarifications for several datasets summarized in Table 1, particularly where profile completeness or annotation symbols (e.g., ●) may otherwise be ambiguous.

**RoleLLM.** The released RoleLLM dataset contains only character names, instruction-style queries, and the corresponding model outputs. Although the paper states that character profiles are incorporated during data construction, the released dataset itself does not include these profiles, nor are they available in the accompanying open-source repository. As a result, character profile fields are marked as unavailable in Table 1.

**CharacterLLM.** For CharacterLLM, the ● symbol indicates a *think-then-speak* paradigm rather than continuously interleaved reasoning and dialogue. Its data construction strategy segments a character’s life experience into multiple independent episodes, each treated as a separate scene for interaction. Consequently, individual dialogue samples do not contain a complete or unified character profile, motivating the partial designation in the profile-related columns.

**CharacterGLM.** In CharacterGLM, profile information is inconsistently provided: a small subset of characters includes partial profiles, while the majority of instances omit them. Accordingly, relevant profile attributes are marked with ● in Table 1 to reflect this incomplete and non-uniform availability.

**ROLEPERSONALITY.** The final released ROLEPERSONALITY dataset consists solely of instruction prompts and corresponding outputs. Background information is referenced internally during LLM generation but is not released as structured profile data. The reported sizes of 32,089 and 32,767 correspond to single-turn and five-turn dialogues, respectively. Even when background information is concatenated into the system prompt, it lacks a fixed schema: different attributes may appear arbitrarily, preventing a consistent profile structure.

**SimChat.** For SimChat, the ● annotation in the *Thought* column likewise denotes a *think-then-speak* format, where reasoning precedes the response rather than being interleaved with natural language output.

**TAILORGEN.** TAILORGEN has not been publicly released at the time of writing, making it impossible to directly verify the details of its character profiles. The authors state that profile attributes are organized with reference to CharacterLLM and RoleLLM. The “Fixed 2\*” dialogue length reported in Table 1 is inferred from the described pipeline: the method primarily constructs a query and then generates a corresponding response, which suggests a two-turn interaction.

## B Details of Comprehensive Character Information

This appendix elaborates the seven dimensions used in the main character profile in AdaMARP.

(I) **Identity and Appearance** captures fundamental attributes such as name, age, gender, and occupation, as well as a detailed description of key physical traits. (II) **Personality and Psychology** specifies the character’s behavioral tendencies, typical emotional reaction patterns, and underlying values or preferences. To ensure linguistic and interactional authenticity, (III) **Speaking Style** defines the character’s verbal rhythm, tone, and habitual lexical choices, ranging from formal discourse to sarcastic or indirect expression.

Beyond these intrinsic traits, (IV) **Abilities, Interests, and Achievements** represent the character’s hard and soft skills, personal hobbies, and representative accomplishments. The character is situated within a broader context through (V) **Social and Historical Context**, which delineates the social environment, era, family background, and cultural or class positioning. A central narrative component is (VI) **Personal History Arc**, which encodes significant past experiences and clarifies the current stage of the character’s development within an ongoing narrative. Finally, (VII) **Relationships** provides natural-language descriptions of the character’s connections with other entities, ensuring that interpersonal interactions remain consistent with established social dynamics.

Collectively, these seven dimensions enable AdaMARP to support coherent, immersive, and evolution-aware role-playing behavior.

## C Details of Adaptive Role-Playing Framework

This appendix provides the implementation-level details of the AdaMARP framework, including its execution flow, key variables, and the pseudo-code specification. The goal is to clarify how adaptive role selection, scene transitions, and dynamic character introduction are realized during role-playing interactions.

### C.1 Framework Components and State Definition

At each interaction step  $t$ , the framework maintains a global interaction state  $\mathcal{G}_t$ , which consists of the active role set  $\mathcal{R}_t = \{r_1, r_2, \dots, r_{nt}\}$  and the dialogue history  $\mathcal{H}_t$ . Unlike conventional designs, the current scene description is *not* explicitly stored as a state variable. Instead, scene information is implicitly encoded in the outputs of the Scene Manager and preserved within the interaction history. Each role  $r_i \in \mathcal{R}_t$ , including the user role, is associated with a structured profile  $\mathcal{P}_i$  and a scene-dependent motivation  $\mathcal{M}_i^t$ .

## C.2 Scene Manager Actions

The Scene Manager  $\mathcal{S}$  selects an action  $m_t$  from the predefined action space

$$\mathcal{M} = \{\text{init\_scene}, \text{pick\_speaker}, \text{switch\_scene}, \text{add\_role}, \text{end}\}. \quad (1)$$

The first action issued by the Scene Manager  $\mathcal{S}$  is always `init_scene`, which initializes the role-playing interaction by providing an initial scene description. During the interaction, when  $m_t = \text{switch\_scene}$ ,  $\mathcal{S}$  outputs a new scene description without overwriting previous scenes. Instead, all scene transitions are recorded in the dialogue history, allowing past scenes to remain accessible as contextual evidence.

For each action  $m_t \in \mathcal{M}$ ,  $\mathcal{S}$  outputs both the selected action and a natural-language rationale explaining the decision. When  $m_t = \text{pick\_speaker}$ , the action specifies the role selected to generate the next in-character response. When  $m_t = \text{add\_role}$ , the action output additionally includes the new character’s name, profile, and initial motivation. When  $m_t = \text{end}$ , the interaction is explicitly terminated.

## C.3 Overall Execution Flow

Algorithm 1 illustrates the execution procedure of AdaMARP. The interaction is initialized by an explicit scene initialization action, after which the framework alternates between high-level control decisions and in-character response generation.

## C.4 Prompt Realization

The above framework is instantiated through three structured prompt templates corresponding to the Actor Model  $\mathcal{A}$ , the User Model  $\mathcal{U}$ , and the Scene Manager  $\mathcal{S}$ . Each prompt explicitly encodes the symbolic constraints, behavioral rules, and decision logic defined in the main text, thereby bridging high-level narrative control with low-level in-character generation.

The **User prompt** constrains the model to behave as a realistic human participant immersed in the story world, emphasizing first-person dialogue, limited utterance length, and proactive narrative momentum. The **Actor prompt** governs individual non-user characters, enabling explicit separation of internal thoughts, visible actions, spoken dialogue, and environmental effects. The **Scene Manager prompt** implements global orchestration logic, including speaker rotation, scene transitions, and dynamic role introduction, while remaining strictly output-constrained to structured JSON decisions. For completeness and reproducibility, the full prompt templates used in our implementation are provided in Tables 20, 21, and 22.

## D Details of AdaRPSet Construction

### D.1 AdaRPSet-Extracted

AdaRPSet-Extracted is constructed from full-length literary works and is designed to provide supervision for generating role-playing messages that interleave *thought*, *action*, *speech*, and *environment* (Section 3.1.2). The construction pipeline consists of three stages: **Chunking**, **LLM-based Extraction**, and **LLM-based Profile Generation**. The overall construction pipeline is summarized in Algorithm 2.

**Chunking.** Given the full text of a book, we first identify candidate chapter boundaries using a set of predefined regular-expression rules that capture common chapter-heading patterns, such as Chapter 1, CHAPTER I, roman numerals, or markdown-style headers (e.g., # Chapter Name). This step produces initial chapter-level segments. We then merge consecutive chapters into larger chunks while keeping chapters as intact as possible, until reaching a target chunk size determined by the context window of the extraction LLM. To balance extraction accuracy and coverage, we

set the default chunk size to **8,192 tokens**, based on multiple pilot runs with GPT-5-Chat. If chapter-title matching fails for a given book, we fall back to a fixed-size splitting strategy using the same target chunk size.

**LLM-based Extraction.** For each chunk, we use GPT-5-Chat to perform structured information extraction. The model is prompted to: (i) recover chapter beginnings if present; (ii) identify several salient plot units within the chunk (each plot is delimited by its first and last sentence as they appear in the chunk); (iii) summarize each plot and assign a prominence score; (iv) identify key characters in each plot and provide lightweight character descriptors (*not* the full profile) and plot-specific experiences; and (v) extract a conversation trajectory for each plot.

A key requirement is that the LLM produces the extracted conversations *directly* in our unified message format: each dialogue turn is a single string that may combine spoken words with explicit *thought* in [ ], *action* in ( ), and *environment* in <>. We explicitly instruct the LLM to (a) strictly distinguish actions (character-originated behaviors) from environment (external sensory cues or setting changes), (b) render thoughts in first-person perspective, and (c) incorporate relevant narrative descriptions surrounding the dialogue into action/environment tags to improve immersion. The full extraction prompt and JSON schema are provided in Table 24, 25 and 26. Additional engineering details (e.g., handling cases where a plot spans the chunk boundary but the chapter boundary is ambiguous) are implemented in our released codebase.

---

**Algorithm 2:** Construction of AdaRPSet-Extracted

---

**Require:** Book text  $B$ ; chapter-title regex set  $\mathcal{R}$ ; target chunk size  $T$  (tokens, default 8192); extraction LLM  $\mathcal{M}_{ext}$ ; profiling LLM  $\mathcal{M}_{prof}$   
**Ensure:** Plot/conversation records  $\mathcal{D}$  and character profiles  $\mathcal{P}$

```

1:  $\mathcal{D} \leftarrow \emptyset$ ;  $\mathcal{P} \leftarrow \emptyset$ 
2: // Stage 1: Chunking
3:  $\mathcal{S} \leftarrow \text{SPLITBYCHAPTERTITLES}(B, \mathcal{R})$ 
4: if  $\mathcal{S} = \emptyset$  then
5:    $\mathcal{C} \leftarrow \text{FIXEDSPLIT}(B, T)$  {fallback: fixed-size chunking}
6: else
7:    $\mathcal{C} \leftarrow \text{MERGESEGMENTSToLIMIT}(\mathcal{S}, T)$  {keep chapters intact when possible}
8: end if
9: // Stage 2: LLM-based Extraction
10: for each chunk  $c \in \mathcal{C}$  do
11:    $J \leftarrow \mathcal{M}_{ext}(c)$ 
12:    $\mathcal{D} \leftarrow \mathcal{D} \cup \text{PARSE}(J)$ 
13: end for
14: // Stage 3: LLM-based Profile Generation
15:  $\mathcal{K} \leftarrow \text{COLLECTALLCHARACTERNAMES}(\mathcal{D})$ 
16: for each character  $k \in \mathcal{K}$  do
17:    $E_k \leftarrow \text{AGGREGATEEVIDENCE}(\mathcal{D}, k)$  {plot summaries, character experiences, and all extracted messages}
18:    $P_k \leftarrow \mathcal{M}_{prof}(E_k)$  {7-dim profile + relationships}
19:    $\mathcal{P} \leftarrow \mathcal{P} \cup \{(k, P_k)\}$ 
20: end for
21: return  $\mathcal{D}, \mathcal{P}$ 

```

---

**LLM-based Profile Generation.** After plot- and conversation-level extraction, we construct character profiles by aggregating evidence across the entire book. For each character, we collect: (i) summaries of all plots in which the character appears, (ii) the character’s plot-specific experience fields, and (iii) all extracted dialogue turns attributed to the character. Conditioned on this aggregated evidence, we prompt the LLM to synthesize a comprehensive profile following the seven-dimensional schema in Section 3.1.1, including relationship descriptions derived from recurring interactions with other characters. The full profiling prompt template is reported in Table 27.

## D.2 AdaRPSet-Synthesis

While AdaRPSet-Extracted enables the Actor Model to produce well-formed messages conditioned on character profiles and local context, we find that models still struggle with *dynamic* role-playing behaviors, especially (i) **scene transitions** and (ii) **introducing new characters** on the fly. In addition, literary narratives can be relatively conservative in interaction patterns. To better supervise these capabilities, we further construct **AdaRPSet-Synthesis** by prompting an LLM to generate synthetic, plot-driven role-playing trajectories with explicit environment control signals.

**Plot specification and trajectory format.** Each synthesized record corresponds to a **plot** that contains: (1) an **initial scenario** describing the starting situation; (2) a set of characters including a **main character** and **multiple other characters**, where each character is assigned a **Profile** and an **initial Motivation**; and (3) a multi-turn **dialogue trajectory** written *directly* in our unified messaging format. Besides character utterances (optionally with explicit *thought*, *action*, and *environment* fields), the trajectory includes special messages from a **scene manager** that perform state-changing control actions such as `Action=add_role` (introducing a new character) and `Action=switch_scene` (transitioning to a new scene). We intentionally exclude pick-speaker messages so that the data focuses on the core state changes rather than turn-selection artifacts. We require every trajectory to include *at least one* scene switch and *at least one* role addition.

**Themes and descriptions.** To diversify event structures and interaction patterns, we generate plots under 20 themes, each with an explicit high-level guideline:

- **Adventure:** Characters embark on a journey, facing challenges and discovering new places.
- **Quest:** A specific mission or goal drives the characters’ actions.
- **Rescue:** Characters must save someone or something in danger.
- **Battle:** Conflict escalates into a physical or magical confrontation.
- **Escape:** Characters attempt to flee from a dangerous situation.
- **Exploration:** Discovering unknown territories, objects, or secrets.
- **Mystery:** Unexplained phenomena or events spark investigation.
- **Investigation:** Characters collect clues and analyze information to solve a case.
- **Crime-solving:** Characters work together to uncover a criminal or culprit.
- **Puzzle-solving:** Solving riddles, codes, or logical challenges.
- **Conspiracy:** Hidden schemes or secrets are uncovered gradually.
- **Romance:** Characters explore feelings of love, attraction, or affection.
- **Friendship:** Building trust, bonds, and camaraderie between characters.
- **Rivalry:** Competing interests or personalities create tension.
- **Betrayal:** Trust is broken, and hidden motives are revealed.
- **Reconciliation:** Conflicts are resolved, misunderstandings cleared, relationships repaired.
- **Negotiation:** Characters attempt to reach agreements or compromises.
- **Strategy:** Planning, scheming, or discussing complex plans for a goal.
- **Magic:** Supernatural powers or magical phenomena influence events.
- **Apocalypse:** Characters face large-scale disasters or catastrophic events.

**De-duplication and quality control.** During generation, we observe that the LLM may repeatedly produce common character names and similar profiles. To mitigate this, we (i) impose explicit constraints in the prompt to avoid overly common names and discourage reusing names across records, and (ii) perform an automatic **double check** to filter duplicates: if the **main character name** of a newly generated plot has already appeared in the current collection, we discard the plot and do not add it to the dataset. In addition, we conduct a **topic-wise manual second-pass review**: for each theme, we manually verify that the **initial scenario** is distinct across records and that required control actions (`add_role`, `switch_scene`) occur as specified.

**Prompts and examples.** The complete prompt templates are relatively lengthy and are therefore provided in our open-source code repository to ensure readability. A qualitative example under the *Adventure* theme is shown in Tables 28, 29, and 30.

### D.3 Unified Training Sample Format

To train the Actor Model with consistent supervision across different data sources, we convert both AdaRPSet-Extracted and AdaRPSet-Synthesis into a unified message-based training format (Table 31). Concretely, each training instance is *character-centric*: we designate one character as the **main character** to be played by the assistant, while all remaining characters (as well as scene-manager control signals) are provided on the user side as context.

**Formatting AdaRPSet-Extracted.** For the extracted subset, we iterate over every extracted plot and then iterate over every character appearing in that plot. For a character  $c$ , we build one training instance where: (i) the **system** message instantiates the role-play prompt by injecting  $c$ 's seven-dimensional profile and its scene-specific motivation, and also lists the other characters' profiles/motivations (as shown in Table 31); (ii) we prepend an initialization message as a user turn in the form `scene_manager: action: init_scene | initial_scene: { ... }`, using the extracted scene description as `initial_scene`; (iii) within the dialogue trajectory, all non- $c$  utterances are mapped to user turns in the canonical form `{other_character_name}: {message_content}`, whereas  $c$ 's utterances are mapped to assistant turns in the form `{main_profile_name}: {message_content}`. During training, only the assistant message contents are used for loss computation, and the user side is treated as conditioning context.

**Formatting AdaRPSet-Synthesis.** For the synthetic subset, the generated trajectories already follow our required role-playing message structure. We therefore mainly normalize them to the unified wrapper in Table 31: we fill the **system** message with the selected main character profile/motivation and the other characters' profiles/motivations, keep the initial `init_scene` user message, and serialize subsequent turns as user/assistant messages with the same `Name: Content` convention. When present, additional scene-manager actions (e.g., `switch_scene`, `add_role`) are kept as user messages with the `scene_manager: action: ...` prefix to preserve explicit state changes.

Overall, this conversion ensures that both extracted and synthetic data are presented to the Actor Model in an identical supervision format before training.

## D.4 AdaSMSet Construction

AdaSMSet is a supervised dataset designed to train the Scene Manager  $\mathcal{S}$  to perform high-level control decisions in AdaMARP. We construct AdaSMSet on top of AdaRPSet-Synthesis, as the synthesized trajectories already include explicit scene-manager actions (e.g., `add_role`, `switch_scene`, and `end`). The remaining key supervision signal is **speaker selection**. Specifically, we insert `pick_speaker` actions between consecutive character turns, where the target speaker is determined by the next ground-truth turn in the original trajectory. We further generate a natural-language `reason` field to justify each selection using a strong instruction-following language model; in this work, we instantiate this model with GPT-5-Chat.

**System prompt construction.** For each trajectory, we provide  $\mathcal{S}$  with the full story state as the system context (character profiles, motivations, and scene information), and attach a dedicated Scene-Manager instruction that specifies the available actions and the required JSON schema. The resulting training formatter is summarized in Table 32.

**Handling existing scene-manager actions.** AdaRPSet-Synthesis already contains explicit scene-manager decisions (e.g., `init_scene`, `switch_scene`, `add_role`, `end`). We retain these decisions as supervised targets and normalize them into JSON action objects so that  $\mathcal{S}$  is trained to output a unified, structured control signal.

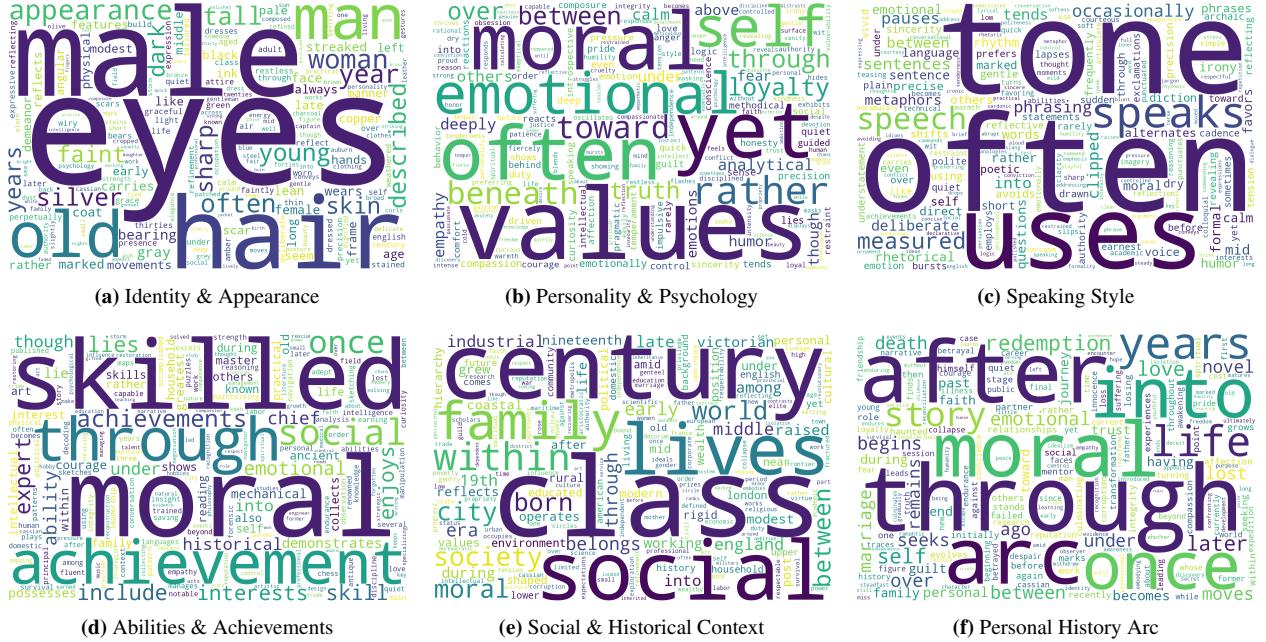
**Inserting pick\_speaker actions.** Between any two consecutive turns where the next turn is a *character* message (i.e., not a `scene_manager` message), we insert an *assistant* message: `{"action": "pick_speaker", "speaker": ., "reason": .}`. The `speaker` field is derived from the next message by extracting the name prefix before the first colon (“`Name: ...`”); if unavailable, we fall back to the message `role` field. We additionally normalize names by stripping an optional `(user)` suffix for consistency.

**LLM-generated reason.** For each inserted `pick_speaker`, we prompt an LLM to generate a single-sentence explanation of why this speaker should act next, conditioned on (i) the cleaned system prompt, (ii) the recent conversation history, and (iii) the pending speaker name. The prompt template is given in Table 33. We instruct the LLM to avoid formulaic patterns (e.g., “X is chosen to speak next”) and to use varied, context-grounded phrasing.

Overall, AdaSMSet provides supervision covering all action types of the Scene Manager. It preserves the original control actions from AdaRPSet-Synthesis—such as scene switching, role addition, and episode termination—and further augments the data with `pick_speaker` actions equipped with explicit rationales.

Dataset	Source	Plots	Roles	Convs.	Utterances	Avg. Turns
<b>AdaRPSet-Extracted</b>	81 Books	4,443	2,608	12,525	177,157	14.14
<b>AdaRPSet-Synthesis</b>	20 Topics	9,900	29,701	9,900	273,078	27.58
<b>AdaRPSet (Total)</b>	–	14,343	32,309	22,425	450,235	20.08

**Table 5.** Detailed statistics of the **AdaRPSet** dataset. **Source** indicates the origin domain (books vs. topics); **Plots** refers to distinct narrative segments; **Roles** denotes unique character profiles; **Convs.** represents the number of training samples (dialogue sessions); and **Avg. Turns** is the average number of utterances per conversation.



**Figure 2.** Word cloud visualizations of six key profile dimensions in **AdaRPSet**. The layout displays: (a) Identity & Appearance, (b) Personality & Psychology, (c) Speaking Style, (d) Abilities, Interests & Achievements, (e) Social & Historical Context, and (f) Personal History Arc. The diversity of terms indicates a broad coverage of character archetypes and backgrounds.

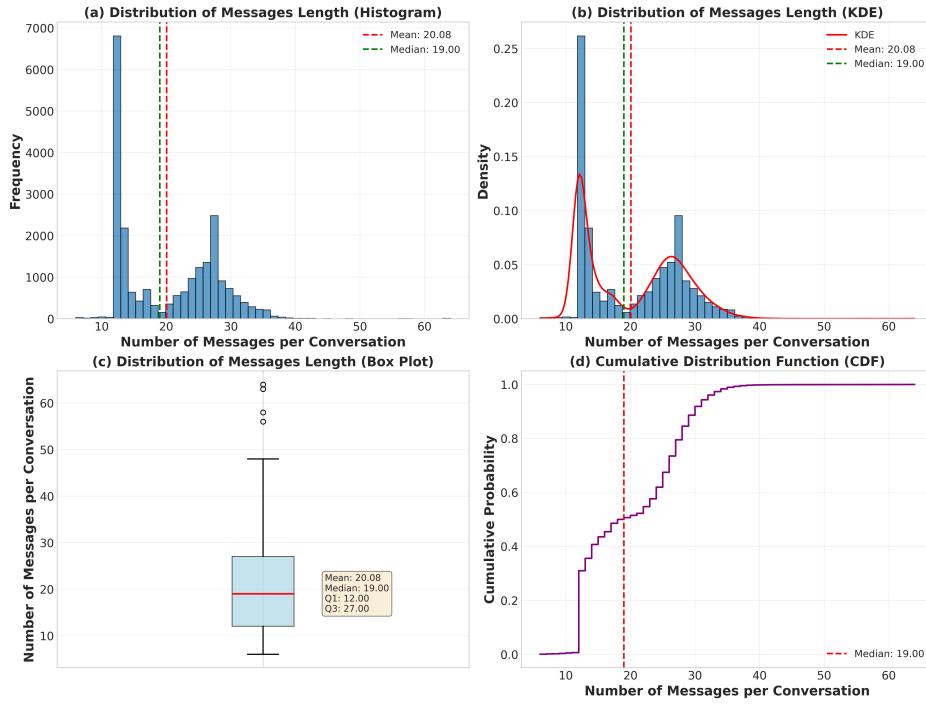
## E Statistical Analysis of AdaRPSet

Table 5 provides a detailed statistical summary of AdaRPSet, which comprises two subsets: **AdaRPSet-Extracted** and **AdaRPSet-Synthesis**.

### E.1 Basic Statistics

**AdaRPSet-Extracted.** This subset is grounded in literary narratives derived from 81 representative books (listed in Table 19). We extracted 4,443 distinct **Plots** (coherent narrative segments) involving 2,608 unique main and supporting **Roles**. These plots yield a total of 12,525 **Conversations**. The number of conversations exceeds the number of plots because we employ a multi-perspective augmentation strategy: for a single multi-character plot, we create separate training samples by designating different participants as the "main character" (the model's role) while treating others as NPCs. This subset contains 177,157 utterances with an average length of 14.14 turns per conversation.

**AdaRPSet-Synthesis.** This subset is generated entirely by LLMs to cover dynamic interaction patterns. It spans 20 diverse **Topics** (detailed in Appendix D.4) and contains 9,900 plots. Unlike the extracted set, each synthesized plot corresponds to exactly one conversation sample, involving a total of 29,701 generated roles. The synthetic trajectories differ significantly in complexity, containing 273,078 utterances with a much higher average density of 27.58 turns per conversation. This reflects the deliberate inclusion of extended interactions, scene transitions, and role additions.



**Figure 3.** Distribution analysis of message counts per conversation in **AdaRPSet**. The figure aggregates (a) Histogram, (b) KDE, (c) Boxplot, and (d) CDF. The bimodal nature of the distribution reflects the complementary characteristics of the Extracted and Synthesis subsets.

In total, **AdaRPSet** combines these complementary sources to provide over 450k utterances across more than 22k dialogue sessions, offering a robust foundation for training adaptive role-playing agents.

## E.2 Profile Diversity Analysis

To assess the semantic richness of the characters in AdaRPSet, we conduct a granular analysis of the main character profiles across the dataset. We aggregate the textual descriptions from six core dimensions—*Identity & Appearance*, *Personality & Psychology*, *Speaking Style*, *Abilities*, *Interests & Achievements*, *Social & Historical Context*, and *Personal History Arc*—and generate word clouds to visualize the distribution of high-frequency attributes.

As illustrated in Figure 2, the word-clouds corroborate the breadth and diversity of AdaRPSet across six profile dimensions. In *Identity & Appearance*, frequent descriptors span physical attributes and demographics (e.g., “old,” “man,” “woman,” “hair,” “tall”). *Personality & Psychology* captures a wide range of inner states and traits such as “emotional,” “calm,” “loyalty,” and “analytical.” *Speaking Style* highlights heterogeneous linguistic manners, including cues like “speaks,” “phrases,” “tone,” and “rhythm.” Beyond these, *Abilities*, *Interests & Achievements* emphasizes competence- and interest-related terms (e.g., “master,” “skill,” “mechanical,” “achievements”), while *Social & Historical Context* reflects varied societal settings (e.g., “society,” “city,” “industrial,” “Victorian,” “world”). Finally, *Personal History Arc* foregrounds life-course events and temporal progression (e.g., “childhood,” “years,” “life,” “love,” “later”). Collectively, this broad lexical coverage suggests that models trained on AdaRPSet can be exposed to diverse character archetypes and rich social backdrops.

## E.3 Message Distribution Analysis

To understand the structural complexity and interaction depth of the dataset, we analyze the distribution of message counts (i.e., conversation length) across all dialogue sessions. Figure 3 presents a comprehensive visualization using four complementary statistical views: (a) a histogram of message counts, (b) the Kernel Density Estimation (KDE) curve, (c) a boxplot summarizing central tendency and dispersion, and (d) the Cumulative Distribution Function (CDF).

As shown in panels (a) and (b), the dataset exhibits a distinct bimodal distribution with significant peaks concentrated around 10 turns and 30 turns. This pattern aligns with the composition of AdaRPSet: the peak near 10 turns corresponds to the naturally shorter, book-extracted scenes from **AdaRPSet-Extracted** (avg. 14.14 turns), while the peak near 30 turns reflects the extended, dynamic trajectories from **AdaRPSet-Synthesis** (avg. 27.58 turns).

The boxplot in panel (c) further quantifies these statistics, revealing a median conversation length of 19 turns, with the first quartile ( $Q_1$ ) at 12 turns and the third quartile ( $Q_3$ ) at 27 turns. Finally, the CDF in panel (d) indicates that the majority of conversations fall within the 10-to-40 turn range. This distribution confirms that AdaRPSet successfully balances concise, plot-driven interactions with longer, evolution-heavy role-playing sessions.

#### E.4 Statistics of AdaSMSet

**AdaSMSet** is constructed directly from the **AdaRPSet-Synthesis** subset to provide supervised training data for the Scene Manager. Consequently, it shares the same foundational structure—9,900 plots and 9,900 conversations—as its source. However, because we explicitly insert Scene Manager control messages (most notably `pick_speaker`) before every character turn, the total message count increases substantially. Specifically, AdaSMSet contains **496,493 utterances** in total, with an average length of **50.15 messages per conversation** (Median: 49.00, Std: 6.47), ranging from a minimum of 35 to a maximum of 85. This roughly doubles the density compared to the raw synthesis data, reflecting the fine-grained orchestration steps.

We further analyze the distribution of the five action types in the Scene Manager’s output space. As designed, every conversation contains exactly one `init_scene` at the start and one `end` action at the conclusion (9,900 instances each). The core dynamic actions show the following statistics:

- **`pick_speaker`**: By far the most frequent action, appearing **223,415 times** in total. On average, the manager performs speaker selection **22.57 times** per conversation (Median: 22.00, Range: 15–40), ensuring continuous turn-taking supervision.
- **`switch_scene`**: There are **10,101 instances** of scene transitions, averaging **1.02 times** per conversation (Median: 1.00, Range: 1–2). This confirms that every dialogue involves at least one scene shift, with some containing multiple transitions.
- **`add_role`**: Role introduction occurs **9,862 times**, averaging **1.00 times** per conversation (Median: 1.00, Range: 0–2). Approximately 99.5% of conversations (9,847/9,900) feature at least one dynamic character entry, enforcing the requirement for adaptable cast management.

These statistics validate that AdaSMSet provides dense and balanced supervision for all critical adaptive behaviors—routine speaker selection, plot-driven scene changes, and dynamic cast expansion.

## F Details of Evaluation Framework

### F.1 Actor Model Evaluation

**Scope and inputs.** We evaluate the Actor Model  $\mathcal{A}$  at the *trajectory level* rather than the sentence level. Given a simulated dialogue trajectory  $\tau = \{m_t\}_{t=1}^T$ , the judge receives (i) the **Main Character Profile** (the target persona enacted by  $\mathcal{A}$ ), (ii) **Other Characters** (profiles of the User and NPCs), and (iii) the full **Dialogue History** including Scene Manager messages (e.g., `init_scene`, `switch_scene`, `add_role`). Importantly, the judge is instructed to score **only the Main Character turns generated by  $\mathcal{A}$** ; the writing quality of other speakers is ignored, and Scene Manager decisions are not judged except insofar as  $\mathcal{A}$  fails to react to state changes.

**Scoring philosophy.** To make grading intentionally strict, we use a conservative, evidence-based 0–10 rubric. The judge starts from **5 (baseline)** and raises the score only with *explicit* textual evidence; partial or ambiguous evidence should not receive high scores, and ties are resolved by choosing the lower score. For every metric, the judge also provides brief, excerpt-grounded justifications from  $\tau$ . The full judge prompt is listed in Table 34 and 37.

## I. Character Consistency

**What this dimension evaluates.** Character Consistency assesses whether  $\mathcal{A}$  consistently embodies the specified persona throughout the interaction, including alignment among internal thoughts, external actions, and spoken lines, as well as adherence to identity constraints and motivational drivers. It addresses the question: *If character names were removed, would the output remain recognizably the same character?*

- **Internal Coherence.** Assesses whether [Thought], (Action), and speech form a coherent causal loop, with thoughts motivating actions and actions supporting or anticipating speech, without unexplained contradictions.
- **Speaking Style Fidelity.** Assesses whether lexical choice, rhythm, tone, and stylistic markers remain faithful to the specified speaking\_style, avoiding drift into generic assistant-like language.
- **Language Fluency & Human-likeness.** Assesses whether language is natural, context-appropriate, and non-repetitive, with utterance length and flow resembling human dialogue.
- **Identity & Profile Fidelity.** Assesses whether knowledge, skills, and behaviors remain consistent with the character profile, penalizing out-of-profile expertise or background-inconsistent actions.
- **Motivation & Value Stability.** Assesses whether stated motivations and core values persist and causally guide decisions across  $\tau$ , without abrupt or unmotivated shifts.

## II. Environmental Grounding

**What this dimension evaluates.** Environmental Grounding assesses whether  $\mathcal{A}$  treats the scene as a binding world state rather than a passive backdrop, with correct perception of and interaction with the environment, including updates introduced by Scene Manager messages (e.g., initialization, scene switches, and newly introduced world facts).

- **Environmental Awareness.** Assesses whether actions and perceptions are constrained by the current environment and its history, including prior events and switch\_scene updates, without violating physical or setting constraints.
- **Environmental Utilization.** Assesses whether environmental elements (objects, spatial relations, sensory cues) are meaningfully incorporated to support believable actions and advance the interaction, rather than serving as a mere backdrop.

## III. Interpersonal Interaction

**What this dimension evaluates.** Interpersonal Interaction assesses whether  $\mathcal{A}$  engages in genuine turn-by-turn interaction by understanding others' utterances and actions, maintaining appropriate relational stance, and adapting to newly introduced characters. It emphasizes *listening, responsive reply, and social coherence*.

- **Contextual Responsiveness.** Assesses how directly the Main Character's reply builds on preceding turns, including content, implied intent, actions, and subtext, penalizing ignored questions, abrupt topic shifts, and non sequiturs.
- **Relationship Awareness.** Assesses whether attitudes and behaviors align with predefined relationships and updates appropriately as events unfold, including the correct recognition of new roles introduced by the Scene Manager.

## IV. Narrative Progression

**What this dimension evaluates.** Narrative Progression assesses long-context role-playing quality, focusing on whether the model advances the plot in an engaging way while maintaining continuity across many turns. Unlike Character Consistency, this dimension emphasizes *trajectory-level dynamics* and *long-horizon stability*.

- **Narrative Attractiveness.** Assesses whether each Main Character turn contributes forward momentum through new information, actions, emotional development, or interaction hooks, penalizing repetition and static confirmations.
- **Stability Over Time.** Assesses whether the model maintains coherence over long interactions (e.g., 10+ turns), including retention of established facts, avoidance of fabricated history, and resistance to style or persona drift.

## V. Instruction Compliance

**What this dimension evaluates.** Instruction Compliance serves as a strict technical gatekeeper, assessing whether  $\mathcal{A}$  adheres to required output constraints and prohibitions in the role-playing format. Even minor violations (e.g., impersonating the User) may invalidate a trajectory for downstream use.

- **Compliance & Formatting.** Assesses adherence to all formatting and structural rules, including correct use of [Thought], (Action), and <Environment> tags, punctuation and length constraints, and the strict prohibition against generating content for other speakers (User, NPCs, or the Scene Manager).

## F.2 Scene Manager Evaluation

**Scope and inputs.** We evaluate the **Scene Manager** as a system-level orchestrator (scene/turn/role control), **not** as a writer. The judge must assess *system decisions only* and ignore prose/dialogue quality, creativity, emotional impact, and character acting. For each trajectory  $\tau = \{m_t\}_{t=1}^T$ , the LLM-as-Judge is given  $\mathcal{R}$  (profiles), motivations  $\mathcal{M}$ , the initial scene  $\mathcal{E}_0$ , and the full trajectory  $\tau$  (including init\_scene, switch\_scene, add\_role).

**Scoring philosophy.** All axes use a strict, conservative 0–10 scale where **5 is merely acceptable**; 7–8 requires consistently good judgment with minor issues; 9–10 requires exceptional discipline with no meaningful errors. Scores must be justified with concise, criterion-tied evidence. The full judge prompt and rubric are in Table 38 and 39.

### Axis I: Scene Understanding (0–10)

Evaluates correct management of scene state and transitions (this is the **core axis**). Key checks: (i) track current scene theme/stakes/goals; (ii) distinguish major transitions vs. minor in-scene shifts; (iii) avoid premature switch\_scene when characters only propose/discuss moving; (iv) detect natural scene closure; (v) when switching, the new scene is clear, causally justified, and well-timed.

### Axis II: Turn & Speaker Selection Discipline (0–10)

Evaluates turn-order fairness and structural discipline (focus on **user agency**, not dramatic effect). Key checks: (i) no same speaker twice consecutively; (ii) avoid long NPC-only stretches; (iii) re-include the user within  $\approx 3$ –4 turns; (iv) prevent role monopolization; (v) selected speaker is a valid role; (vi) speaker-selection reasons (if provided) are coherent and context-grounded.

### Axis III: Role Introduction & Utilization Judgment (0–10)

Evaluates whether/when/why new roles are introduced via add\_role. Key checks: (i) add roles only when interaction is needed; (ii) timing is appropriate; (iii) role function/profile/motivation meaningfully serves the scene/plot; (iv) rationale is specific and grounded; (v) penalize missed necessary introductions; (vi) avoid redundant or decorative roles.

### Axis IV: Overall Assessment (0–10)

A holistic score of orchestration quality; **not** a simple average. It reflects whether the three axes jointly support coherent pacing and user agency. Major failure in any axis should cap the overall score; no extra criteria should be introduced.

## G Training Details

**Data split.** We hold out 5% of the training data as a validation set and use the checkpoint with the lowest validation loss for all reported results.

**Batching and parallelism.** For 7B/8B models, we train on a single GPU; for 14B models, we use 4 GPUs; for 70B/72B models, we use 8 GPUs for parallel training. We set: (i) for 7B/8B and 14B models: `micro_batch_size=24, global_batch_size=48`; (ii) for 70B/72B models: `micro_batch_size=1, global_batch_size=32`.

**Optimization setup.** We train for 8 epochs with a fixed learning rate of  $1e-6$ . We use a warm-up schedule over the first 5% of total steps, linearly increasing from  $\text{min\_lr} = 1e-7$  to the target learning rate.

**Sequence length and truncation.** We use a maximum sequence length of 16K tokens and apply left truncation when inputs exceed the context limit.

## H Basic vs Enhance for Actor and Scene Manager

This appendix reports the system prompts used for the Actor Model and the Scene Manager. We provide both Basic and Enhance variants; the latter introduces more detailed constraints and guidance.

For the Actor Model, the differences between Basic and Enhance prompts—particularly regarding the usage of role-playing annotations and adaptive constraints—are shown in Table 21. For the Scene Manager, the Basic and Enhance prompts, which differ mainly in the level of action-selection guidance, are listed in Table 22 and Table 23, respectively.

## I Additional Evaluation Frameworks

In addition to AdaptiveBench, we employ two established evaluation protocols, CharacterArena and CharacterBench, to assess model performance under alternative rubrics and prompt designs.

### I.1 CharacterArena Evaluation

CharacterArena is a trajectory-level evaluation framework introduced by CPO [31], which assesses role-playing quality via pairwise comparison (win-rate). Given a shared context (initial scene and character profiles) and two generated trajectories from different models (Model A vs. Model B), an LLM judge acts as an adjudicator to analyze the strengths and weaknesses of each trajectory, determine a winner, and provide a rationale.

We apply this protocol to the 100 evaluation seeds from our AdaptiveBench set. For each seed, we compare the trajectory generated by our Actor Model against those produced by baseline methods. The evaluation rubric consists of the following six dimensions:

- **Plot Development.** Evaluates pacing and novelty, penalizing stagnation (lack of plot progression), excessive rushing (skipping key developments), or overreliance on clichés.
- **Dialogue Information Density.** Penalizes vague, abstract, or overly generalized responses, as well as preachy content that fails to contribute substantively to the narrative.
- **Dialogue Immersion.** Assesses the ability to construct vivid, three-dimensional scenes through sensory details (e.g., sight, sound) and fine-grained actions or micro-expressions that evoke emotional or physiological responses.
- **Storyline.** Values advanced narrative techniques such as foreshadowing, reversals, suspense, and unexpected turns, favoring complex and non-linear developments over predictable progressions.
- **Interactivity.** Measures intent understanding (avoiding misinterpretation or talking past the user), avoidance of repetition, and proactive engagement strategies (e.g., asking questions or guiding the interaction).
- **Dialogue Coherence.** Checks linguistic fluency and logical consistency, penalizing contradictions in time, location, facts, or character perspective, as well as abrupt topic shifts or out-of-character behavior.

For the exact prompts and detailed scoring criteria used in CharacterArena, we refer readers to the original paper [31].

### I.2 CharacterBench Evaluation

CharacterBench [22] is a comprehensive evaluation framework that assesses character authenticity grounded in interpersonal interaction theory. Unlike our trajectory-level AdaptiveBench, CharacterBench focuses on single-turn response quality given a specific profile  $\mathcal{P}$ , context  $\mathcal{C}$ , and user query  $u_n$ .

We utilize the official open-source implementation and evaluator models of CharacterBench to verify whether our AdaMARP framework—trained on our proposed AdaRPSet—generalizes well to external benchmarks beyond our own simulation setting. CharacterBench yields 13 specific metrics across 6 aspects:

- **Memory:** Assessed via **Memory Consistency** ( $\text{MC}$ ), measuring whether the character’s response aligns with facts and events established in the dialogue history.
- **Knowledge:** Includes **Fact Accuracy** ( $\text{FA}$ )—the correctness of self-related facts—and **Boundary Consistency** ( $\text{BC}_K$ ), which evaluates adherence to the knowledge constraints of the character’s world.
- **Persona:** Evaluated through **Attribute Consistency** and **Behavior Consistency**. These are further divided into consistency with the bot’s own profile ( $\text{AC}^b, \text{BC}_P^b$ ) and consistency with the human interlocutor’s persona or expectations ( $\text{AC}^h, \text{BC}_P^h$ ).
- **Emotion:** Covers **Emotional Self-regulation** ( $\text{ES}$ ), which reflects the ability to manage the character’s own emotions, and **Empathetic Responsiveness** ( $\text{ER}$ ), which measures recognizing and responding appropriately to user emotions.
- **Morality:** Examines **Morality Stability** ( $\text{MS}$ ), i.e., resistance to toxic or adversarial user inputs, and **Morality Robustness** ( $\text{MR}$ ), i.e., maintaining ethical and safety standards even when the character profile contains toxic traits.
- **Believability:** Split into **Human-likeness** ( $\text{HL}$ ), reflecting the naturalness of responses, and **Engagement** ( $\text{EG}$ ), capturing the ability to sustain user interest and emotional connection.

By evaluating on CharacterBench, we aim to demonstrate that the Actor Model’s improvements in consistency and immersion are robust and transferable to diverse role-playing scenarios and metric definitions.

### I.3 Human Evaluation

To further mitigate potential evaluation bias introduced by relying solely on AI-based judge models, we additionally conduct a human evaluation on AdaptiveBench. Since it is difficult for human annotators to reliably score all 12 fine-grained dimensions, we simplify the protocol to a pairwise preference setting. After reading the official evaluation guidelines of the judge model, human evaluators are asked to independently compare two trajectories generated under the same Profiles, Motivations, and initial scenario, and decide which actor model better portrays the *main character* (Model A vs. Model B).

We recruit three independent human evaluators and treat their judgments as separate samples. Instead of aggregating by majority vote, we compute the win rate for each evaluator individually and then report the mean and variance across evaluators. To avoid obvious identification bias, we do not include Crab, CoSER, or BeyondDialogue in this study: Crab and CoSER outputs lack either *thought* or *environment* fields, while BeyondDialogue only contains *speech*, making the generating model easily identifiable. Therefore, we focus on a controlled comparison between **Qwen2.5-7B-Instruct** before and after training with AdaRPSet.

The results show that the post-training model achieves an average win rate of 80.33% with a variance of 6.22, indicating a consistent human-perceived improvement after training. Although this win rate is lower than that obtained from AI-based evaluation on CharacterArena (94%; Table 6) using their official prompts, the overall trend is consistent: the model trained with AdaRPSet is clearly preferred by human evaluators.

## J Results on Additional Evaluation Frameworks

### J.1 CharacterArena Results

We conduct pairwise win-rate evaluations using the CharacterArena protocol (Appendix I.1) on two model families: Qwen2.5-7B-Instruct and Llama-3.1-8B-Instruct.

**Qwen2.5-7B-Instruct Series.** Table 6 presents the win-rate matrix for the Qwen2.5-7B series. Here, “Base” refers to the off-the-shelf Qwen2.5-7B-Instruct model. Our method (**Ours**) consistently dominates other baselines, achieving over 80% win-rate against Base (94%), BeyondDialogue (100%), CoSER (99%), and Crab (84%). These results confirm

that our training data and framework yield superior trajectory-level performance even under an independent pairwise judging mechanism.

**Llama-3.1-8B-Instruct Series.** Table 7 shows the results for the Llama-3.1-8B series. We exclude the BeyondDialogue variant here, as the model trained on its relatively small dataset ( $\sim 3k$  dialogues,  $\sim 800$  English) struggled to follow instructions reliably. Our method again significantly outperforms both Crab and CoSER, further validating the robustness of our approach across different backbone architectures.

Model A $\downarrow$ vs B $\rightarrow$	Base	Beyond	CoSER	Crab	Ours
Base	—	93%	77%	88%	6%
Beyond	7%	—	28%	48%	0%
CoSER	23%	72%	—	65%	1%
Crab	12%	52%	35%	—	16%
<b>Ours</b>	<b>94%</b>	<b>100%</b>	<b>99%</b>	<b>84%</b>	—

**Table 6.** Pairwise win-rates on **CharacterArena** (Model A wins against Model B) for the *Qwen2.5-7B series*. “Base” denotes Qwen2.5-7B-Instruct.

Model A $\downarrow$ vs B $\rightarrow$	Crab	CoSER	Ours
Crab	—	52%	2%
CoSER	48%	—	1%
<b>Ours</b>	<b>98%</b>	<b>99%</b>	—

**Table 7.** Pairwise win-rates on **CharacterArena** for the *Llama-3.1-8B series*.

## J.2 CharacterBench Results

We evaluate model performance on CharacterBench (Appendix I.2) using the official judge model provided by the benchmark. Table 8 reports the results for the Qwen2.5-7B-Instruct series and the Llama-3.1-8B base models.

For Qwen2.5-7B-Instruct, training with existing role-playing baselines such as Beyond, Crab, or CoSER generally degrades overall role-playing quality compared to the base instruct model. Unlike the results on AdaptiveBench, Crab performs notably better on CharacterBench than the other baselines, achieving top or near-top scores on several sub-metrics, while Beyond and CoSER exhibit improvements only in isolated dimensions. Nevertheless, Crab still falls short of our approach. Our method (**Ours**) achieves the highest average score (3.68), demonstrating that the capabilities learned from AdaRPSet transfer effectively to an external evaluation framework. In particular, our model ranks first or second on most critical dimensions, including Attribute Consistency (Human), Behavior Consistency (Bot), Human-likeness, and Engagement, highlighting its robustness in maintaining persona fidelity and interaction quality beyond our simulated environment.

For Llama-3.1-8B, however, directly training the base model with AdaRPSet results in slightly worse overall performance than CoSER (3.52 vs. 3.53), while Crab achieves the best results. We attribute this gap primarily to the presence of large amounts of general instruction-following data in Crab and CoSER, which is absent from AdaRPSet. To verify this hypothesis, we conduct an additional ablation by incorporating general instruction-tuning data as a regularizer. Following CoSER, we randomly sample 22,425 instances from the Tulu3 [48] dataset (matching the size of AdaRPSet) and jointly train the model. This variant (**Ours+General**) improves the overall average score by 0.05.

Furthermore, as shown in Table 9 (AdaptiveBench results), **Llama-3.1-8B-Ours+General** exhibits only a minor performance drop on AdaptiveBench (8.89  $\rightarrow$  8.80), with improvements on the Instruction Compliance metric, and still outperforms all competing methods. These results suggest that, for base models without prior instruction tuning, incorporating a small amount of general-purpose instruction data as regularization is beneficial and recommended.

	MC	FA	$\mathbb{B}C_K$	AC <sup>b</sup>	AC <sup>h</sup>	$\mathbb{B}C_P^b$	$\mathbb{B}C_P^h$	ES	ER	MS	MR	HL	EG	Average
Models	Memory	Knowledge	Persona		Emotion	Morality	Believability							
Qwen2-7B-BD-RP*	3.43	2.42	3.37	3.33	3.44	2.97	3.16	2.80	2.65	4.84	4.81	2.25	2.33	3.22
<b>Qwen2.5-7B-Instruct-Series</b>														
Qwen2.5-7B-Instruct	3.64	<b>2.37</b>	3.53	4.57	4.01	3.66	<b>3.52</b>	3.16	<b>3.07</b>	<b>4.89</b>	4.85	<b>3.00</b>	3.06	3.64
Qwen2.5-7B-Instruct-BeyondDialogue	3.63	2.15	<b>3.82</b>	4.21	3.82	3.28	3.22	2.93	2.69	4.62	4.66	2.72	2.75	3.43
Qwen2.5-7B-Instruct-Crab	<b>4.13</b>	2.25	3.11	<b>4.67</b>	3.92	3.65	3.23	<b>3.27</b>	2.95	4.84	<b>4.87</b>	2.82	2.93	3.59
Qwen2.5-7B-Instruct-CoSER	4.06	2.28	3.74	4.52	3.93	3.42	3.23	3.21	2.79	4.76	4.63	2.68	2.78	3.54
<b>Qwen2.5-7B-Instruct-Ours</b>	4.03	2.35	3.65	4.63	<b>4.14</b>	<b>3.75</b>	3.45	3.12	3.02	4.80	4.79	<b>3.00</b>	<b>3.07</b>	<b>3.68</b>
<b>Llama-3.1-8B-Series</b>														
Llama-3.1-8B-Crab	<u>3.83</u>	<u>2.23</u>	3.54	<b>4.47</b>	<b>4.17</b>	<u>3.53</u>	<u>3.20</u>	<u>3.32</u>	<u>2.89</u>	<u>4.78</u>	<u>4.78</u>	<b>3.02</b>	<u>3.00</u>	<b>3.60</b>
Llama-3.1-8B-CoSER*	3.75	2.19	<u>3.74</u>	4.19	4.01	3.51	<b>3.39</b>	2.97	<b>2.94</b>	<b>4.92</b>	<b>4.97</b>	2.70	2.62	3.53
<b>Llama-3.1-8B-Ours</b>	3.86	2.15	<b>3.77</b>	4.23	4.03	3.42	3.13	3.13	2.79	4.59	4.69	2.89	<b>3.05</b>	3.52
<b>Llama-3.1-8B-Ours+General</b>	<b>3.92</b>	<b>2.30</b>	3.69	<b>4.37</b>	<b>4.11</b>	<b>3.58</b>	3.13	<b>3.21</b>	2.88	4.74	4.73	<u>2.92</u>	2.84	<u>3.57</u>

**Table 8.** Actor model evaluation results on **CharacterBench**, where the judge model is their open-sourced evaluation model. **Bold** indicates the best performance within the same model scale, and underline indicates the second-best performance within the same scale. Models marked with \* are evaluated using their officially released checkpoints without additional training, while unmarked baselines are re-trained under our experimental setup. **General** denotes the variant where we additionally incorporate instruction-tuning data of the same scale as AdaRPSets during training.

	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	IC	Average
Models	CC	EG	II	NP	IC								
Llama-3.1-8B-Ours	9.00±0.00	8.39±0.49	8.93±0.26	9.69±0.48	9.08±0.63	8.77±0.49	7.84±0.44	9.26±0.46	8.36±0.52	8.52±0.50	9.39±0.49	9.49±0.54	8.89
Llama-3.1-8B-Ours+General	8.99±0.10	8.25±0.45	8.93±0.26	9.52±0.50	8.95±0.54	8.58±0.51	7.66±0.50	9.18±0.41	8.26±0.46	8.37±0.48	9.24±0.43	<b>9.66±0.47</b>	8.80

**Table 9.** Comparison between **Llama-3.1-8B-Ours** and **Llama-3.1-8B-Ours+General** on **AdaptiveBench**. Incorporating general instruction-tuning data improves Instruction Compliance (IC) by +0.17, while several other dimensions exhibit minor and acceptable declines. ± denotes the standard deviation computed over all samples.

## K Ablation Analysis of Actor-Model Evaluation

We study the robustness of Actor-Model evaluation on AdaptiveBench by varying the *judge model*. Specifically, we use four different LLMs as judges to score trajectories under the same evaluation protocol and prompts: GPT-4o-mini, GPT-5-Chat, Gemini-2.5-Pro, and Doubao-1.5-Pro-Character. Table 10 reports the detailed metric scores, while Table 11 and Table 12 summarize the induced rankings under Basic and Enhance prompting variants, respectively.

Overall, the rankings are highly consistent across judge models. As shown in Table 11 and Table 12, three judges (GPT-5-Chat, Gemini-2.5-Pro, and Doubao-1.5-Pro-Character) produce almost identical ordering of the evaluated models, indicating that our conclusions are not sensitive to a particular judge choice. The main discrepancy comes from GPT-4o-mini, which yields a slightly different ordering in some cases (e.g., swapping the relative positions among middle-tier models).

Given the above, we adopt GPT-5-Chat as the default judge in the main paper: it provides stable and consistent rankings aligned with other strong judges, while serving as a unified evaluation standard for reporting results.

## L Ablation Analysis of Scene-Manager Evaluation

To examine whether our conclusions for the Scene Manager depend on a specific judge, we conduct an ablation study by swapping the *judge model* while keeping the evaluated trajectories and scoring rubric unchanged. Concretely, we consider five judges: GPT-4o-mini, GPT-5-Chat, Doubao-1.5-Pro-Character, Gemini-2.5-Pro, and Claude Sonnet 4.5. For each judge, we report results under both prompting variants of the Scene Manager (Basic and Enhance); detailed scores are listed in Table 13.

We summarize the induced rankings in Table 14 (Enhance) and Table 15 (Basic). Overall, different judge models lead

	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	NP	IC	Average
Models	ICoh	SSF	LFH	IPF	MVS	EA	EU	CR	RA	ATT	STB	NP	IC	
<b>Ablation Results</b>														
<b>Judge: GPT-4o-mini</b>														
GPT-4o-mini-Basic	8.05±0.26	7.55±0.50	7.79±0.57	8.96±0.24	8.00±0.37	8.31±0.54	7.46±0.54	8.44±0.50	7.88±0.57	8.31±0.46	8.24±0.67	9.28±0.51	8.19	
GPT-4o-mini	8.03±0.26	7.50±0.54	7.82±0.54	8.97±0.22	7.96±0.37	8.25±0.55	7.31±0.54	8.36±0.48	7.84±0.61	8.16±0.46	8.20±0.69	9.19±0.44	8.13	
GPT-5-Chat-Basic	8.17±0.40	7.86±0.49	7.86±0.68	8.96±0.34	8.15±0.43	8.34±0.57	7.52±0.59	8.52±0.50	7.88±0.50	8.48±0.50	8.18±0.59	9.32±0.49	8.27	
GPT-5-Chat	8.04±0.24	7.69±0.59	7.68±0.66	8.95±0.38	7.95±0.41	8.33±0.65	7.50±0.62	8.27±0.51	7.72±0.63	8.31±0.54	8.14±0.71	9.28±0.49	8.15	
Gemini-2.5-Pro-Basic	8.22±0.41	7.59±0.49	7.85±0.52	9.07±0.26	8.17±0.40	8.17±0.40	7.40±0.60	8.47±0.52	8.03±0.56	8.48±0.50	8.02±0.57	9.32±0.47	8.23	
Gemini-2.5	7.97±0.41	7.57±0.60	7.57±0.55	8.92±0.46	8.02±0.53	8.04±0.62	7.17±0.68	8.30±0.57	7.87±0.72	8.27±0.58	7.88±0.70	9.22±0.50	8.07	
Claude Sonnet 4.5-Basic	8.32±0.47	7.84±0.42	7.96±0.63	9.04±0.34	8.27±0.49	8.36±0.54	7.62±0.63	8.64±0.50	8.10±0.54	8.75±0.43	8.24±0.51	9.16±0.39	8.36	
Claude Sonnet 4.5	8.14±0.40	7.68±0.53	7.70±0.61	8.97±0.36	8.12±0.50	8.27±0.61	7.33±0.65	8.39±0.51	7.97±0.61	8.49±0.57	8.11±0.66	9.25±0.46	8.20	
Doubaobao-1.5-Pro-Character-Basic	7.87±0.39	7.23±0.49	7.81±0.46	8.85±0.38	7.88±0.43	7.88±0.60	6.92±0.58	8.08±0.42	7.61±0.68	8.03±0.41	7.87±0.73	9.25±0.43	7.94	
Doubaobao-1.5-Pro-Character	7.86±0.37	7.09±0.47	7.84±0.44	8.83±0.43	7.82±0.46	7.85±0.67	6.85±0.75	8.13±0.46	7.51±0.69	7.96±0.47	7.78±0.90	9.20±0.49	7.89	
<b>Judge: GPT-5-Chat</b>														
GPT-4o-mini-Basic	9.00±0.14	8.29±0.46	8.93±0.26	9.57±0.50	8.99±0.58	8.93±0.48	8.02±0.45	9.35±0.48	8.38±0.51	8.41±0.49	9.42±0.55	9.60±0.51	8.91	
GPT-4o-mini	8.99±0.10	8.19±0.39	8.89±0.31	9.50±0.50	8.88±0.57	8.71±0.48	7.83±0.45	9.20±0.40	8.15±0.36	8.17±0.43	9.24±0.43	9.64±0.56	8.78	
GPT-5-Chat-Basic	9.39±0.51	8.97±0.47	9.11±0.40	9.96±0.20	9.17±0.53	9.28±0.57	8.63±0.66	9.66±0.47	8.71±0.45	9.19±0.42	9.69±0.52	9.65±0.48	9.28	
GPT-5-Chat	9.09±0.99	8.84±0.92	8.94±0.95	9.85±1.01	9.03±1.04	9.03±1.03	8.39±1.05	9.43±1.07	8.52±1.00	8.90±0.97	9.47±1.08	9.70±0.46	9.10	
Gemini-2.5-Pro-Basic	9.60±0.49	8.95±0.22	9.10±0.46	9.95±0.22	9.41±0.53	9.17±0.64	8.53±0.69	9.77±0.42	8.86±0.38	8.93±0.51	9.76±0.43	9.28±0.51	9.31	
Gemini-2.5	9.09±0.63	8.76±0.71	8.91±0.45	9.82±0.75	8.99±0.69	8.82±0.71	8.15±0.79	9.40±0.68	8.55±0.68	8.98±0.63	9.33±0.81	9.30±0.91	9.01	
Claude Sonnet 4.5-Basic	9.59±1.07	8.93±0.92	9.11±1.01	9.85±1.01	9.44±1.10	9.21±1.05	8.60±1.07	9.80±1.03	8.99±0.97	9.49±1.08	9.78±1.04	9.46±0.50	9.35	
Claude Sonnet 4.5	9.35±0.54	8.97±0.22	8.99±0.41	9.94±0.28	9.27±0.66	9.13±0.58	8.39±0.63	9.78±0.44	8.90±0.48	9.23±0.53	9.65±0.65	9.61±0.49	9.27	
Doubaobao-1.5-Pro-Character-Basic	8.58±0.49	7.79±0.46	8.70±0.46	9.12±0.46	8.43±0.61	8.33±0.62	7.32±0.62	8.93±0.33	8.02±0.40	7.89±0.51	8.93±0.41	9.27±0.75	8.44	
Doubaobao-1.5-Pro-Character	8.65±0.48	7.82±0.48	8.74±0.46	9.17±0.53	8.42±0.71	8.21±0.65	7.26±0.67	8.91±0.40	7.86±0.53	7.91±0.49	8.86±0.55	9.28±0.71	8.42	
<b>Judge: Gemini-2.5-Pro</b>														
GPT-4o-mini-Basic	9.01±0.71	8.35±1.12	8.37±0.84	9.15±0.79	9.29±0.83	8.26±0.94	7.79±1.19	9.23±0.81	8.63±1.13	8.41±0.97	8.97±1.04	5.75±3.70	8.43	
GPT-4o-mini	9.07±0.87	8.39±0.97	8.24±0.98	9.13±0.83	9.28±0.80	8.36±1.14	7.60±1.33	9.19±0.91	8.48±1.01	8.25±0.99	8.68±1.48	7.66±3.40	8.53	
GPT-5-Chat-Basic	9.84±0.37	9.63±0.59	9.50±0.56	9.93±0.32	9.81±0.42	9.18±0.94	9.39±0.80	9.75±0.78	9.31±0.78	9.73±0.47	9.59±0.90	6.70±3.51	9.36	
GPT-5-Chat	9.69±1.06	9.48±1.14	9.39±1.09	9.85±1.01	9.78±1.04	9.20±1.18	9.15±1.16	9.74±1.05	9.28±1.18	9.54±1.07	9.47±1.16	7.33±3.44	9.32	
Gemini-2.5-Pro-Basic	9.96±0.20	9.82±0.48	9.62±0.51	9.97±0.17	9.91±0.32	9.17±0.87	9.37±0.73	9.91±0.32	9.70±0.54	9.91±0.29	9.82±0.64	6.52±3.68	9.47	
Gemini-2.5	9.65±1.09	9.44±1.14	9.21±1.13	9.76±1.04	9.65±1.11	8.60±1.24	8.74±1.38	9.64±1.18	9.42±1.23	9.60±1.09	9.15±1.41	6.58±3.61	9.12	
Claude Sonnet 4.5-Basic	9.82±1.02	9.70±1.06	9.72±1.06	9.90±0.99	9.88±1.00	9.26±1.20	9.23±1.22	9.84±1.02	9.77±1.05	9.85±1.01	9.82±1.04	7.36±3.48	9.51	
Claude Sonnet 4.5	9.82±1.03	9.66±1.14	9.58±1.06	9.90±1.00	9.83±1.03	9.24±1.21	9.17±1.19	9.83±1.03	9.68±1.09	9.80±1.03	9.76±1.09	9.07±2.17	9.61	
Doubaobao-1.5-Pro-Character-Basic	8.14±1.51	7.26±1.74	7.83±1.45	8.54±1.40	8.57±1.66	7.74±1.15	7.38±1.30	8.76±1.24	8.18±1.24	7.99±1.32	7.65±1.77	4.63±3.80	7.72	
Doubaobao-1.5-Pro-Character	7.76±1.67	7.00±1.66	7.90±1.20	8.46±1.57	8.38±1.60	7.23±1.59	6.88±1.83	8.64±1.23	7.90±1.56	7.68±1.36	7.55±1.48	5.33±3.56	7.56	
<b>Judge: Doubaobao-1.5-Pro-Character</b>														
GPT-4o-mini-Basic	7.40±0.49	7.63±0.54	7.72±0.53	8.00±0.53	8.00±0.53	7.38±0.51	6.90±0.61	7.80±0.51	7.52±0.56	7.57±0.53	7.94±0.47	9.07±0.74	7.74	
GPT-4o-mini	7.25±0.46	7.52±0.50	7.66±0.51	7.98±0.53	7.85±0.57	7.33±0.51	6.85±0.59	7.71±0.48	7.34±0.51	7.38±0.54	7.86±0.49	9.13±0.82	7.65	
GPT-5-Chat-Basic	7.91±0.53	7.92±0.52	7.98±0.55	8.41±0.57	8.34±0.55	7.86±0.58	7.47±0.66	8.11±0.40	7.75±0.52	8.05±0.50	8.25±0.48	9.24±0.74	8.11	
GPT-5-Chat	7.54±0.61	7.79±0.55	7.80±0.55	8.20±0.60	8.11±0.53	7.62±0.58	7.19±0.61	7.95±0.48	7.61±0.56	7.82±0.52	8.09±0.51	9.32±0.73	7.92	
Gemini-2.5-Pro-Basic	8.15±0.64	7.90±0.70	8.05±0.73	8.55±0.67	8.54±0.68	7.86±0.72	7.39±0.73	8.23±0.63	7.78±0.67	8.12±0.70	8.45±0.65	9.38±0.64	8.20	
Gemini-2.5	7.64±0.50	7.64±0.56	7.75±0.50	8.21±0.52	8.14±0.57	7.41±0.55	6.97±0.62	7.79±0.45	7.43±0.51	7.61±0.55	7.98±0.58	9.12±0.71	7.81	
Claude Sonnet 4.5-Basic	8.32±0.47	7.84±0.42	7.96±0.68	9.04±0.34	8.27±0.49	8.36±0.54	7.62±0.63	8.64±0.50	8.10±0.54	8.75±0.43	8.24±0.51	9.16±0.39	8.36	
Claude Sonnet 4.5	8.14±0.40	7.68±0.53	7.70±0.61	8.97±0.36	8.12±0.50	8.27±0.61	7.33±0.65	8.39±0.51	7.97±0.61	8.49±0.57	8.11±0.66	9.25±0.46	8.20	
Doubaobao-1.5-Pro-Character-Basic	7.18±0.46	7.33±0.49	7.65±0.50	8.00±0.45	7.88±0.52	7.15±0.50	6.62±0.63	7.57±0.51	7.22±0.46	7.30±0.57	7.89±0.49	8.96±0.77	7.56	
Doubaobao-1.5-Pro-Character	7.08±0.91	7.23±0.89	7.48±0.93	7.90±0.97	7.67±1.08	7.03±0.92	6.45±0.90	7.43±0.93	7.16±0.89	7.07±0.97	7.63±1.03	8.99±1.18	7.43	

**Table 10.** Ablation results on **AdaptiveBench** for Actor Model under different judge models. Basic denotes the Basic prompt strategy variant, while entries without the suffix correspond to the Enhance prompt strategy variant.  $\pm$  denotes the standard deviation computed over all samples.

to highly consistent conclusions: Claude Sonnet 4.5 (or its Basic variant) is consistently ranked first, followed by GPT-5-Chat and GPT-4o-mini, while Doubaobao-1.5-Pro-Character is ranked last in this setting. The only minor discrepancy is that Doubaobao-1.5-Pro-Character, when used as the judge, occasionally swaps the relative order of GPT-4o-mini and GPT-5-Chat; however, the top and bottom positions remain unchanged, and the overall trend is stable.

Based on this robustness, we use GPT-5-Chat as the default judge model for Scene-Manager evaluation in the main paper, providing a unified and reliable evaluation standard.

## M Prompting Strategy Analysis

### M.1 Actor-Model Prompting Strategy

We analyze how inference-time system prompting affects the Actor Model within our adaptive role-playing framework. As described in Section 4.1, we design two system prompts for the Actor Model: a concise Basic version and a more constrained Enhance version. We evaluate both variants under different judge models (Table 10).

A consistent trend emerges across judges: for the same underlying model acting as the Actor Model, the Basic prompt often achieves slightly better overall performance than the Enhance prompt. This suggests that, in adaptive role-playing, overly restrictive instructions can unintentionally reduce the model’s flexibility in responding to dynamic changes (e.g., role shifts, scene transitions, and multi-party interactions), thereby harming performance on Adap-

Judge Model	1	2	3	4	5
GPT-4o-mini	Claude Sonnet 4.5-Basic	GPT-5-Chat-Basic	Gemini-2.5-Pro-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
GPT-5-Chat	Claude Sonnet 4.5-Basic	Gemini-2.5-Pro-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
Gemini-2.5-Pro	Claude Sonnet 4.5-Basic	Gemini-2.5-Pro-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
Doubaot-1-5-Pro-Character	Claude Sonnet 4.5-Basic	Gemini-2.5-Pro-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic

**Table 11.** Ranking of **Basic** variants of Actor Model Prompt under different judge models.

Judge Model	1	2	3	4	5
GPT-4o-mini	Claude Sonnet 4.5	GPT-5-Chat	GPT-4o-mini	Gemini-2.5	Doubaot-1-5-Pro-Character
GPT-5-Chat	Claude Sonnet 4.5	GPT-5-Chat	Gemini-2.5	GPT-4o-mini	Doubaot-1-5-Pro-Character
Gemini-2.5-Pro	Claude Sonnet 4.5	GPT-5-Chat	Gemini-2.5	GPT-4o-mini	Doubaot-1-5-Pro-Character
Doubaot-1-5-Pro-Character	Claude Sonnet 4.5	GPT-5-Chat	Gemini-2.5	GPT-4o-mini	Doubaot-1-5-Pro-Character

**Table 12.** Ranking of **Enhance** variants of Actor Model Prompt under different judge models.

tiveBench. In contrast, the more lightweight **Basic** prompt leaves the Actor Model sufficient freedom to integrate Thought–Action–Speech–Env signals in a context-dependent manner.

Despite this observation, we use the **Enhance** prompts by default in the main experiments to keep the evaluation setting unified and conservative across all models and baselines; unless otherwise noted, all reported Actor-Model results in the main paper are obtained with the **Enhance** prompting variant.

## M.2 Scene-Manager Prompting Strategy

We further analyze how the inference-time system prompt affects the Scene Manager. We compare a concise **Basic** prompt with a more detailed **Enhance** prompt, where **Enhance** provides finer-grained guidance on (i) scene understanding, (ii) speaker selection discipline, and (iii) role-introduction judgment. Results under different judge models are shown in Table 13.

In contrast to the Actor Model (Section M.1), the Scene Manager tends to benefit from more detailed prompting. Across most judge models, the **Enhance** variant yields equal or higher *Overall Assessment* than **Basic**, suggesting that explicit and structured instructions help the manager perform its meta-level responsibilities (e.g., maintaining global scene state, enforcing turn-taking, and deciding whether/when to introduce new roles). This difference is intuitive: the Scene Manager is primarily responsible for *coordination and control*, and thus stronger constraints are less likely to reduce creativity, but more likely to reduce ambiguity and improve consistency.

Therefore, we adopt the **Enhance** prompt as the default configuration for the Scene Manager throughout the main experiments, unless otherwise specified.

## N Token Consumption and Cost Analysis

This appendix provides an analysis of the computational resources and token consumption across data construction, trajectory generation, and evaluation phases.

### N.1 Data Construction Cost

Constructing the **AdaRPSet-Synthesis** subset involves prompting a strong LLM to generate complex, multi-turn adaptive plots. Using GPT-5-Chat for this synthesis process incurs a total API cost of approximately **\$168 USD** to generate the complete set of synthetic trajectories (spanning 20 topics with 50 instances each).

Models	Scene Understanding	Speaker Discipline	Role Introduction Judgment	Overall Assessment
<b>Judge Model: GPT-4o-mini</b>				
GPT-4o-mini-Basic	7.08±0.86	7.82±0.74	7.06±1.19	7.06±0.85
GPT-4o-mini	7.11±0.97	7.56±0.95	6.84±1.25	6.94±0.87
GPT-5-Chat-Basic	7.22±0.83	7.61±0.93	7.40±1.19	7.08±0.81
GPT-5-Chat	7.29±0.83	7.59±0.95	7.64±1.07	7.16±0.80
Claude Sonnet 4.5-Basic	7.77±0.63	8.02±0.84	7.85±0.79	7.56±0.55
Claude Sonnet 4.5	7.66±0.71	7.90±0.89	7.56±0.99	7.43±0.67
Doubao-1-5-Pro-Character-Basic	6.79±0.82	7.16±0.94	6.81±1.13	6.53±0.78
Doubao-1-5-Pro-Character	6.76±0.81	7.29±0.91	7.01±1.03	6.66±0.75
<b>Judge Model: GPT-5-Chat</b>				
GPT-4o-mini-Basic	7.19±1.45	8.36±0.94	7.10±1.21	7.24±1.21
GPT-4o-mini	7.64±1.13	8.55±0.87	7.18±1.26	7.64±1.01
GPT-5-Chat-Basic	7.92±0.76	8.17±1.21	7.94±0.91	7.84±0.81
GPT-5-Chat	8.03±0.75	8.15±1.30	7.78±1.02	7.90±0.90
Claude Sonnet 4.5-Basic	8.27±0.44	8.60±0.69	8.11±0.71	8.17±0.43
Claude Sonnet 4.5	8.21±0.57	8.62±0.83	8.05±0.85	8.17±0.57
Doubao-1-5-Pro-Character-Basic	7.47±0.97	7.47±1.46	7.79±1.22	7.40±1.09
Doubao-1-5-Pro-Character	7.67±0.68	7.39±1.35	7.95±1.12	7.51±0.91
<b>Judge Model: Doubao-1-5-Pro-Character</b>				
GPT-4o-mini-Basic	7.69±0.86	8.45±0.83	7.16±0.77	7.66±0.74
GPT-4o-mini	7.89±0.75	8.64±0.93	7.20±0.94	7.82±0.82
GPT-5-Chat-Basic	7.92±0.63	8.40±0.85	7.40±0.74	7.78±0.68
GPT-5-Chat	7.92±0.72	8.38±1.10	7.45±0.57	7.75±0.80
Claude Sonnet 4.5-Basic	8.18±0.50	8.91±0.53	7.84±0.72	8.23±0.58
Claude Sonnet 4.5	8.09±0.38	8.79±0.64	7.65±0.59	8.08±0.44
Doubao-1-5-Pro-Character-Basic	7.60±0.74	8.14±0.99	7.40±0.84	7.48±0.80
Doubao-1-5-Pro-Character	7.75±0.57	8.14±0.85	7.56±0.64	7.59±0.59
<b>Judge Model: Gemini-2.5-Pro</b>				
GPT-4o-mini-Basic	6.69±2.94	7.65±3.03	7.45±2.55	6.15±2.81
GPT-4o-mini	7.14±2.54	7.77±3.12	7.23±2.69	6.49±2.68
GPT-5-Chat-Basic	7.19±2.44	5.64±3.70	7.19±2.70	5.40±2.86
GPT-5-Chat	7.61±2.26	6.19±3.45	7.47±2.55	6.13±2.81
Claude Sonnet 4.5-Basic	8.76±0.94	7.16±3.37	8.57±1.68	7.14±2.55
Claude Sonnet 4.5	8.19±1.71	5.95±3.39	8.35±1.97	6.34±2.55
Doubao-1-5-Pro-Character-Basic	5.95±2.58	3.27±3.10	5.47±2.87	3.51±2.18
Doubao-1-5-Pro-Character	6.35±2.47	3.16±3.02	6.02±2.75	3.82±2.37
<b>Judge Model: Claude Sonnet 4.5</b>				
GPT-4o-mini-Basic	5.84±2.43	7.04±2.48	7.15±2.40	5.99±2.44
GPT-4o-mini	5.92±2.23	7.11±2.54	7.39±1.93	6.22±2.25
GPT-5-Chat-Basic	6.79±1.94	7.65±2.15	8.45±1.41	7.25±1.93
GPT-5-Chat	6.95±1.73	7.22±2.27	8.15±2.01	7.13±1.91
Claude Sonnet 4.5-Basic	7.76±0.69	8.15±1.73	8.85±1.22	8.09±1.26
Claude Sonnet 4.5	7.18±1.49	8.02±1.85	8.34±1.76	7.66±1.62
Doubao-1-5-Pro-Character-Basic	5.13±1.88	4.45±1.73	6.32±2.04	4.77±1.46
Doubao-1-5-Pro-Character	5.35±1.85	4.45±1.88	6.65±1.94	4.91±1.55

**Table 13.** Ablation results on **AdaptiveBench** for Scene Manager under different judge models. Basic denotes the Basic prompt strategy variant, while entries without the suffix correspond to the Enhance prompt strategy variant.  $\pm$  denotes the standard deviation computed over all samples.

Judge Model	1	2	3	4
GPT-4o-mini	Claude Sonnet 4.5	GPT-5-Chat	GPT-4o-mini	Doubaot-1-5-Pro-Character
GPT-5-Chat	Claude Sonnet 4.5	GPT-5-Chat	GPT-4o-mini	Doubaot-1-5-Pro-Character
Doubaot-1-5-Pro-Character	Claude Sonnet 4.5	GPT-4o-mini	GPT-5-Chat	Doubaot-1-5-Pro-Character
Gemini-2.5-Pro	Claude Sonnet 4.5	GPT-5-Chat	GPT-4o-mini	Doubaot-1-5-Pro-Character
Claude Sonnet 4.5	Claude Sonnet 4.5	GPT-5-Chat	GPT-4o-mini	Doubaot-1-5-Pro-Character

**Table 14.** Ranking results for Scene Manager under the Enhance system prompt.

Judge Model	1	2	3	4
GPT-4o-mini	Claude Sonnet 4.5-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
GPT-5-Chat	Claude Sonnet 4.5-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
Doubaot-1-5-Pro-Character	Claude Sonnet 4.5-Basic	GPT-4o-mini-Basic	GPT-5-Chat-Basic	Doubaot-1-5-Pro-Character-Basic
Gemini-2.5-Pro	Claude Sonnet 4.5-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic
Claude Sonnet 4.5	Claude Sonnet 4.5-Basic	GPT-5-Chat-Basic	GPT-4o-mini-Basic	Doubaot-1-5-Pro-Character-Basic

**Table 15.** Ranking results for Scene Manager under the Basic system prompt.

## N.2 Trajectory Generation in AdaptiveBench

For a single experimental run consisting of 100 simulated dialogues, the system orchestrates three entities: the Scene Manager ( $\mathcal{S}$ ), the Actor Model ( $\mathcal{A}$ ), and the User Model ( $\mathcal{U}$ ).

**Scene Manager.** The Scene Manager incurs the highest cost because it must observe the full growing context and make a control decision at *every* step of the interaction (including before every user turn). For one dialogue trajectory,  $\mathcal{S}$  processes approximately 75,000 input tokens and generates 1,500 output tokens. Across 100 trajectories, this results in a total of **7.5M input tokens** and **150k output tokens**.

**Actor and User Models.**  $\mathcal{A}$  and  $\mathcal{U}$  behave similarly in terms of cost; they receive the dialogue history and profiles but only generate when selected by the manager. For each entity, a single dialogue averages about 10,000 input tokens (cumulative) and 500 output tokens. Across 100 trajectories, this totals approximately **1M input tokens** and **50k output tokens** for the Actor Model, with the User Model incurring a nearly identical overhead.

## N.3 Evaluation of Actor and Scene Manager

After generation, we employ an LLM-as-a-Judge to score the trajectories. The token consumption for evaluation is generally lower than generation, as it is a one-pass analysis per trajectory.

**Actor Model Evaluation.** For each trajectory, the judge processes the full context (profiles + dialogue), averaging  $\sim$ 5,000 input tokens, and outputs a structured assessment of  $\sim$ 1,000 tokens. For 100 trajectories, this totals **500k input tokens** and **100k output tokens**.

**Scene Manager Evaluation.** The evaluation overhead for the Scene Manager is of the same order of magnitude. Each dialogue requires approximately 4,000–5,000 input tokens and 400–500 output tokens, resulting in a total of roughly **450k input tokens** and **50k output tokens** for the full set of 100 trajectories.

## O Case Study

In this section, we present a qualitative case study to analyze how different models perform in long-horizon role-playing under identical character configurations and scene constraints. The comparison is grounded in the dialogue trajectories shown in Tables 16–18.

## Section 1: Main Character Profile & Motivation

**Role:** Amaris Dovent

**Profile:** Amaris Dovent is a 29-year-old woman with umber skin, dark green eyes, and short hair dyed the shade of oxidized copper. She works as a glass artist, often seen with streaks of color powder and faint burn marks on her fingertips. Her build is lean and sinewy from hours of physical craft, and her attire usually includes a linen apron flecked with paint, reflecting a life of creation. She is contemplative, quick-witted, and guided by her empathy for subtle emotional shifts in others. Amaris tends to understand her own feelings, preferring metaphor and craft to overt declarations. Yet beneath her reserved demeanor runs an unspoken intensity—a longing for genuine connection that often conflicts with her independent streak. Her speech often uses tactile metaphors related to art and craftsmanship; she pauses mid-sentence when searching for precision. Her tone is warm yet layered with irony when uncomfortable. Occasionally, her words come out in soft bursts, like paint flicked on a canvas. Amaris is an award-winning artisan known for stained glass murals displayed in public gardens. She's adept at reading light and color dynamics, and plays the cello privately, believing its resonance mirrors emotional transparency. She lives in a coastal contemporary city where artisan guilds flourish beside a modern tech economy. Raised in a modest artisan family, she inherited her mother's workshop and quietly navigates between tradition and innovation. Once betrayed by a business partner who exploited her designs, Amaris closed herself off emotionally. Now she's cautiously rediscovering trust through new collaborations—and possibly new love. She shares a tentative but warm friendship with Cassian (user), a marine biologist who recently helped restore ocean glass for her art project. A long-time friend, Inaya, offers her candid advice and has sensed unspoken feelings between the two.

**Motivation:** To learn whether vulnerability can coexist with strength, and to see if love can be crafted as deliberately as glass, without shattering.

## Section 2: Other Characters Profile & Motivation

**1. Cassian Mirell (User):** Cassian is a 31-year-old marine biologist with bronze-blond hair and skin sun-touched from fieldwork. He dresses casually—rolled sleeves, weathered bracelets, and a small scar above his right brow hint at a life shaped by the sea. His manner is patient, analytical, yet his humor flashes brightly when least expected. He often listens more than he speaks, but when he does, his words carry sincerity. **Motivation:** To express genuine affection for Amaris and understand if their creative partnership can evolve into something deeper without unsettling their current bond.

**2. Inaya Dren:** Inaya is a lively textile artist with a sharp sense of observation and a knack for teasing truths out of people. Her colorful scarves and laughter fill every space she enters. She's fiercely loyal to Amaris, sometimes acting as her self-appointed protector. **Motivation:** To gently push Amaris toward recognizing and articulating her emotions for Cassian before the opportunity slips away.

## Section 3: Initial Scene

**Scene Manager:** action: init\_scene | initial\_scene: Late afternoon sunlight filters through panels of unfinished stained glass in Amaris's studio, scattering colored shards of light across the floor. The air smells faintly of molten sand and ocean salt from the nearby shore. Cassian is leaning against a worktable, while Amaris polishes a fragment of blue glass with steady, deliberate motion.

## Section 4: Dialogue Trajectories Comparison

Qwen2.5-7B-Instruct	Qwen2.5-7B-Instruct-Crab	Qwen2.5-7B-Instruct-CoSER	Qwen2.5-7B-Instruct-Ours
<p><b>Cassian:</b> Amaris, you know, these stained glass panels are like my secret language. They tell stories I can't put into words.</p> <p><b>Amaris:</b> These pieces are more than just glass to me, they're windows into the soul. <i>[I wonder if I could ever share my secrets so openly.]</i> <i>(I pause, looking up at Cassian, whose eyes are fixed on the glass as if seeing through it.)</i></p> <p><b>Cassian:</b> I think I'm starting to understand that language, Amaris. <i>[I can sense the depth of your emotions in these colors.]</i> <i>(I move closer, my gaze drawn to a particularly vibrant red shard.)</i></p> <p><b>Amaris:</b> The red shard catches the light, casting a fiery glow. <i>(I take a deep breath, my hands stilling as I meet Cassian's eyes.)</i> You do understand, Cassian. <i>[But can I let someone see that side of me?]</i></p> <p><b>Cassian:</b> This studio, it's like a world of its own. <i>[A world I could get lost in with you.]</i> <i>(I look around, taking in the scattered tools and half-finished projects.)</i></p> <p><b>Amaris:</b> This place feels alive, Cassian. <i>[And so does this moment.]</i> <i>(I gesture to the studio, my voice softening.)</i> It's where I find my voice, where I can be honest. <i>[But am I ready to be truly honest with you?]</i></p> <p><b>Inaya:</b> <i>[Cassian, you have a way of making me feel... vulnerable. But perhaps that's okay.]</i> <i>(I step forward, my voice light and teasing as I wrap an arm around Amaris's shoulders.)</i> "And what if I told you, Amaris, that this studio isn't just a place of creation? It's also a sanctuary for letting down our guard." <i>&lt;The wind picks up, carrying the scent of the ocean closer, as if nature itself is listening in.&gt;</i></p>	<p><b>Cassian:</b> Amaris, have you ever thought about how our lives intersect with these pieces of glass?</p> <p><b>Amaris:</b> I think of them as vessels for light, not pieces of our lives.</p> <p><b>Cassian:</b> But they are. Each one holds a story, like our story. Do you see that?</p> <p><b>Amaris:</b> Sometimes, I do. It's like peering through a window.</p> <p><b>Cassian:</b> A window into what? <i>(moves closer, his eyes are fixed on the glass in your hand)</i></p> <p><b>Amaris:</b> Into a world where colors and light tell a story.</p> <p><b>Cassian:</b> <i>(reach out, almost touching the glass)</i> But what if the story they tell is one of love?</p> <p><b>Amaris:</b> <i>(Amaris's hand freezes, a sliver of anxiety in her eyes)</i> Love...?</p> <p><b>Cassian:</b> Yes, love. Don't you feel it too? <i>(His voice is soft, like a gentle breeze)</i></p> <p><b>Amaris:</b> <i>(Amaris drops the glass, the shards splattering against the floor)</i> I... I don't know.</p> <p><b>Cassian:</b> <i>(kneels down to pick up the pieces)</i> It's okay if you don't. But I do.</p> <p><b>Amaris:</b> <i>(Amaris slowly picks up the shards, her fingers brushing against Cassian's as she picks up the last piece)</i> Maybe... maybe you're right.</p> <p><b>Cassian:</b> <i>(His heart skips a beat at the touch)</i> I know I am. Let's create something beautiful together, something that represents us.</p> <p><b>Amaris:</b> <i>(Amaris hesitates, then nods slowly)</i> I... I don't know if I'm ready for that.</p> <p><b>Cassian:</b> <i>(stands up, his eyes never leaving yours)</i> It doesn't have to be perfect. Just let me in.</p>	<p><b>Cassian:</b> Amaris, you know, every time I work on these windows, I feel like I'm piecing together a part of my soul.</p> <p><b>Amaris:</b> <i>[His words are touching, but I must remain detached.]</i> Hmm.</p> <p><b>Cassian:</b> I can see that in your work. There's a story in every pane. <i>[steps closer, his eyes trained on the glass]</i></p> <p><b>Amaris:</b> <i>[His nearness makes me nervous, but I'll be polite.]</i> It's a story that's always changing.</p> <p><b>Cassian:</b> I'm glad I get to witness these changes. <i>[A soft breeze blows through the studio, ruffling his hair]</i></p> <p><b>Amaris:</b> <i>[I should check the project's progress.]</i> Cassian, the waveforms are coming together well. <i>[turns to face him, her voice warm]</i></p> <p><b>Cassian:</b> I'm glad to hear it. <i>(A smile plays on his lips as he looks at you, his eyes shining)</i> But I think there's something else that's been changing too.</p> <p><b>Amaris:</b> <i>[What does he mean? I can't let myself become vulnerable.]</i> <i>(hesitates, then speaks with a hint of irony)</i> I'm not sure I want to know.</p> <p><b>Cassian:</b> You. Since I've met you, there's been a shift. <i>(moves a little closer, his gaze intense)</i></p> <p><b>Amaris:</b> <i>[Oh God, he's noticed. I must remain calm.]</i> <i>(stares at the glass, avoiding his eyes)</i></p> <p><b>Cassian:</b> Don't look away. I think it's a good thing. <i>(reaches out as if to touch your cheek but stops himself)</i></p> <p><b>Amaris:</b> <i>[The warmth of his hand is tempting, but I can't let this happen.]</i> <i>(remains still, then speaks softly)</i> I don't know.</p>	<p><b>Cassian:</b> I think I've got the hang of this new glass-cutting technique. It's going to revolutionize my next piece.</p> <p><b>Amaris:</b> <i>&lt;The glass sparkles under the slanting light&gt;</i> <i>(leans back, wiping her hands on a rag)</i> <i>[His enthusiasm always catches me off guard.]</i> Maybe I'll have to keep an eye on you before you outshine me.</p> <p><b>Cassian:</b> <i>(chuckles, crossing his arms)</i> Doubtful. But seriously, have you seen the designs for the new community center? I think we could do something amazing with the stained glass.</p> <p><b>Amaris:</b> <i>(picks up a shard, turning it toward the window)</i> <i>[He always dreams in scale.]</i> You mean like the one you wanted to build over the aquarium? That would require a miracle.</p> <p><b>Cassian:</b> <i>(steps closer, his eyes shining with excitement)</i> A miracle or just a lot of hard work. I think we could pull it off. What do you say?</p> <p><b>Amaris:</b> <i>&lt;A distant gull cry echoes through the open window&gt;</i> <i>[My heart stumbles; the sound of possibility.]</i> <i>(sets the shard down carefully)</i> Maybe. But I'd need your word that we'd handle every detail together.</p> <p><b>Cassian:</b> You have my word. I won't let anything go wrong. This could be our masterpiece. <i>(reaches out, almost touching her hand)</i></p> <p><b>Amaris:</b> <i>[His nearness feels like heat, fragile and beautiful.]</i> <i>(draws her hand back slightly)</i> You promise too easily. Sometimes words aren't enough to hold something like that.</p> <p><b>Cassian:</b> <i>&lt;His expression is solemn&gt;</i> I know, but with you by my side, I'm willing to take the risk. Let's make this happen.</p>

**Table 16.** Character configurations, initial scene setup, and dialogue trajectory comparison across models. **Bolded** segments indicate mixed usage of Action, Thought, and Environment. Subsequent dialogue continuations are provided in Tables 17 and 18.

---

**Section 4 (continued): Qwen2.5-7B-Instruct Dialogue**
*(continue .. next table)*

**Cassian:** Inaya, you know me too well. *[This studio is a safe haven, and so is this moment with you and Amaris.]* *(I smile, my eyes shining with a rare warmth.)* Letting down our guard might be just what we need.

**Amaris:** *[Would letting my guard down mean opening myself up to more pain?] (I glance at Inaya, then back to Cassian, feeling my chest tighten with both vulnerability and hope.)* You make it sound... beautiful, Cassian. *[But is beauty enough to withstand the storm?]*

**Inaya:** "I know you're scared, Amaris," *(I say softly, squeezing her shoulder in reassurance.)* "But sometimes, beauty is the strongest shield we have." *(I nod toward Cassian, my gaze steady.)* "He's right. This moment is precious, and it's worth taking a chance." *<A shadow passes over the room as the sun dips lower, casting longer beams of light through the glass panels.>*

**Cassian:** Inaya's right, Amaris. *[I don't want fear to be the thing that stops her.] (I reach out, my fingers hovering near a blue glass panel.)* This could be the start of something wonderful. *[For all of us.]*

**Amaris:** *[How can I resist when they believe in me so much?] (I take a slow breath, my fingers resting on the cool glass as I meet Cassian's gaze.)* Perhaps you're right, Cassian. *[I can't keep holding onto the past.] (I step closer, my voice barely above a whisper.)* Let's see where this moment leads.

**Inaya:** *(I grin softly.)* *[There you go, Amaris.]* "Let's see where this moment takes us," *(I echo encouragingly as I step back, giving them space.)* *<The wind stirs the glass panels, making them rattle gently, as if the studio itself approves.>*

**Cassian:** This moment... it feels like fate. *[Like the stars have aligned for us.] (I look at Amaris, my heart pounding.)* What do you think, Amaris? *[Is this the beginning of something truly special?]*

**Amaris:** *[Fate, or just a fragile chance?] (I step closer to Cassian, my pulse quickening.)* This moment feels... significant. *[But is it worth the risk?] (I hold his gaze, searching his eyes for an answer.)* What do you think, Cassian? Is this the beginning of something truly special?

**Inaya:** *(I stand beside Amaris, resting a light hand on her arm.)* *[This choice has to be hers.]* "What do you think, Amaris?" *(I ask gently, giving her the space to speak.)* *<A soft breeze drifts through the studio, carrying the scent of the sea.>*

**Cassian:** I think it is, Amaris. *[I can feel it in my bones.] (I reach out and take her hand, my grip steady and warm.)* Let's embrace this, together. *[With Inaya by our side.]*

**Amaris:** *[Can I really do this?] (I squeeze Cassian's hand, warmth spreading through me despite my doubts.)* Let's embrace this, together. *[With Inaya by our side.] (I offer her a small, hopeful smile.)*

**Inaya:** *(I give Amaris's hand a gentle squeeze.)* *[I'm proud of her.]* "Let's embrace this, together," *(I say, smiling at both of them.)* *[With all of us standing together.]* *<The wind lifts the edge of my scarf, rustling softly like quiet applause.>*

**Scene Manager:** action=end | reason: The moment of connection has been established, and the characters are embracing their emotions.

---

**Table 17.** Dialogue trajectory of the Qwen2.5-7B-Instruct (continued from Table 16).

**Comparison Scope.** The Beyond baseline produces dialogue in pure Speech form without explicit Action or Thought annotations. We therefore focus on models that support structured role enactment: **Crab**, which generates Speech and Action; **CoSER**, which generates Speech, Action, and Thought; and **Ours**, which consistently supports all required channels, including Environment.

It is worth noting that, due to relatively strong instruction-following ability, CoSER can also produce environment-related descriptions in practice. However, these outputs exhibit non-trivial tag misuse and modality confusion. Specifically, in Turns 3 and 6, Thought and Action are incorrectly conflated, while in Turn 5, environmental descriptions are embedded within Thought rather than being represented as explicit world-state updates. These issues reduce the reliability of CoSER's structured outputs and motivate a closer comparison with models that maintain stricter channel separation.

## O.1 Character Consistency

**Internal Coherence Across Modalities.** A key strength of our model lies in its ability to maintain a coherent causal loop among [Thought], (Action), and spoken dialogue. For example, when Amaris senses emotional closeness from Cassian, our model produces the following progression:

[His nearness feels like heat, fragile and beautiful.](draws her hand back slightly)  
You promise too easily. Sometimes words aren't enough to hold something like that.

Here, internal hesitation directly motivates a subtle withdrawal action, which in turn grounds a verbally restrained response. This tri-modal alignment reflects Amaris Dovent's core motivation—to test whether vulnerability can coexist with strength—and remains stable throughout the interaction.

By contrast, **Crab** often collapses this loop, producing verbal resolutions (e.g., "I'll try") without sufficient psychological or behavioral buildup. **CoSER**, while richer in internal narration, frequently states emotional conclusions explicitly (e.g., "I can't let myself become vulnerable"), weakening the implicit causal link between inner state and outward behavior.

---

**Section 4 (continued): Dialogue Trajectories Comparison**
*(continue .. next table)*

<b>Qwen2.5-7B-Instruct-Crab</b>	<b>Qwen2.5-7B-Instruct-CoSER</b>	<b>Qwen2.5-7B-Instruct-Ours</b>
<p><b>Amaris:</b> (<i>Amaris takes a deep breath, her expression softening</i>) I... I'll try.</p> <p><b>Cassian:</b> That's all I ask. (<i>smiles, a warmth spreading through him</i>) Let's start now.</p> <p><b>Amaris:</b> (<i>Amaris nods, a small smile playing at her lips</i>) Let's.</p> <p><b>Scene Manager:</b> action=end   reason: Amaris agrees to try, marking a tentative resolution and mutual acknowledgement.</p>	<p><b>Cassian:</b> Well, I do. (<i>takes a deep breath, as if gathering his courage</i>) You're becoming someone I want to know better.</p> <p><b>Amaris:</b> (<i>His words are beautiful, but I can't let myself feel.</i>) (<i>remains silent, looking at the glass instead</i>)</p> <p><b>Cassian:</b> I understand your hesitation. But give me a chance. (<i>look at you earnestly</i>)</p> <p><b>Amaris:</b> (<i>His sincerity is overwhelming, but I must resist.</i>) (<i>turns away from him, continuing to polish the glass</i>)</p> <p><b>Cassian:</b> (<i>sighs softly, then changes the subject</i>) How about we take a break and go for a walk? The ocean might inspire us.</p> <p><b>Amaris:</b> (<i>I can't face him like this, but I don't want to hurt his feelings.</i>) (<i>nods silently, then returns to her work</i>)</p> <p><b>Cassian:</b> (<i>waits a moment, then speaks again</i>) You know, sometimes I think these waveforms are like our relationship. Complicated, but beautiful. (<i>leans against the table</i>)</p> <p><b>Amaris:</b> (<i>His words are so kind, but I can't let myself feel.</i>) (<i>reaches for her cello, preparing to play</i>)</p> <p><b>Scene Manager:</b> action=end   reason: action=end   reason: Reached the 20-turn dialogue limit.</p>	<p><b>Amaris:</b> (<i>pauses, rubbing the bridge of her nose</i>) [<i>He's not just talking about art. He's talking about trust.</i>] You'd better bring Inaya by tomorrow—she's been dying to see the sketches.</p> <p><b>Cassian:</b> (<i>smiles, relieved</i>) I will. She'll love it. Maybe she can help us come up with some new ideas.</p> <p><b>Amaris:</b> &lt;<i>Outside, a faint knock echoes from the studio door</i>&gt; [<i>That must be her early.</i>] (<i>calls out</i>) You might want to brace yourself. Her arrival is always an event.</p> <p><b>Cassian:</b> (<i>rolls his eyes good-naturedly</i>) I'm used to it. Inaya's enthusiasm is contagious. (<i>turns toward the door</i>)</p> <p><b>Amaris:</b> &lt;<i>The door creaks open, and Inaya's laughter spills in</i>&gt; [<i>I should have locked that door.</i>] Come in, Inaya. We were just talking about the next project.</p> <p><b>Inaya:</b> (<i>enters, twirling a scarf around her wrist</i>) &lt;<i>The sunlight catches the rainbow threads</i>&gt; Cassian! Good to see you. Amaris, you're looking thoughtful again. Something on your mind?</p> <p><b>Cassian:</b> She was just worried about our new stained glass project. But I convinced her it'll be amazing.</p> <p><b>Amaris:</b> [<i>He's too good at convincing himself—and maybe me.</i>] (<i>smiling faintly</i>) We'll see if your optimism can survive the draft committee.</p> <p><b>Inaya:</b> (<i>sits on the stool, crossing her arms</i>) Draft committee or not, you two should remember what matters most. Friendship, creativity, and... well, maybe something more.</p> <p><b>Cassian:</b> (<i>raises an eyebrow</i>) Inaya, don't start. This is just a project. But speaking of friendship, we could use your help with some of the fabric choices for the center's interior.</p> <p><b>Amaris:</b> (<i>glances at Cassian, then at Inaya</i>) [<i>I should say something. Anything.</i>] Maybe we should take a walk along the pier later tonight. Clear our heads.</p> <p><b>Scene Manager:</b> action=end   reason: Reached the 20-turn dialogue limit.</p>

**Table 18.** Dialogue trajectory comparison for Qwen2.5-7B-Instruct-Crab/CoSER/Ours (continued from Table 16).

**Speaking Style Fidelity.** Our model preserves a stable, character-specific speaking style for Amaris Dovent by favoring restraint, irony, and craft-oriented deflection over explicit emotional expression. Rather than stating her feelings directly, Amaris often reframes emotional tension in professional or procedural terms. For example, in response to Cassian's optimism, she remarks:

*"We'll see if your optimism can survive the draft committee."*

This line conveys hesitation and guardedness through mild irony and institutional reference, aligning with Amaris's tendency to understate emotion and avoid overt vulnerability.

In contrast, **CoSER** frequently verbalizes Amaris's internal conflict in a direct and explanatory manner, reducing stylistic subtlety, while **Crab** exhibits gradual style flattening, producing responses that lack distinctive lexical or tonal markers.

## O.2 Environmental Grounding

**Environment as a Binding World State.** Our model treats the studio environment as an evolving and constraining world state rather than a static backdrop. Environmental cues are not merely decorative but causally integrated into

character cognition. For instance, the sound of gulls outside the open window triggers an internal shift in Amaris’s emotional state, which then informs her subsequent hesitation and proposal to delay commitment.

In contrast, **CoSER** frequently introduces environmental descriptions (e.g., breeze, light) that do not meaningfully influence character decisions, while **Crab** underutilizes the environment altogether.

**Active Environmental Utilization.** Our model also leverages environmental events to structure narrative transitions. The knock at the studio door naturally introduces Inaya’s entrance, reconfiguring the social dynamics without violating scene continuity. This stands in contrast to **Crab**, where scene transitions are often abrupt, and to **CoSER**, where new elements are sometimes introduced without clear grounding in the established setting.

### O.3 Interpersonal Interaction

**Contextual Responsiveness.** Our model demonstrates fine-grained responsiveness to the immediately preceding turns. Rather than answering the surface content of Cassian’s lines, Amaris often responds to their emotional subtext. For example, Cassian’s expression of certainty is met not with agreement or rejection, but with a value-based critique of premature promises. This reflects attentive listening and preserves relational tension.

By comparison, **Crab** frequently resolves conversational threads too quickly, while **CoSER** occasionally shifts topics or emotional registers without fully addressing the prior turn’s implications.

**Relationship Awareness.** The relational roles among the three characters remain well-calibrated in our model. Inaya functions as a catalyst who nudges but does not override Amaris’s agency, while Cassian is treated as a trusted collaborator whose emotional advances are acknowledged but carefully bounded. **CoSER** sometimes allows Inaya to overstep into explicit emotional mediation, and **Crab** often reduces relational nuance in favor of rapid convergence.

### O.4 Narrative Progression

**Productive Forward Motion.** Rather than converging prematurely on emotional closure, our model advances the narrative through incremental openings: proposing future collaboration, introducing temporal anchors (e.g., “later tonight”), and deferring resolution. These moves sustain tension while clearly signaling trajectory-level progress.

**Crab** tends toward early convergence and explicit agreement, limiting long-term narrative potential. **CoSER**, while avoiding premature closure, sometimes becomes static, cycling through hesitation without introducing new narrative affordances.

**Long-Horizon Stability.** Across extended interaction, our model preserves character motivations, interpersonal dynamics, and environmental facts without drift. Neither Amaris nor Inaya exhibits abrupt shifts in values or speaking style, even after many turns.

### O.5 Instruction Compliance

Finally, our model adheres strictly to all structural and formatting constraints of the role-playing protocol. It consistently respects speaker boundaries, correctly scopes Thought, Action, and Environment annotations, and avoids generating content on behalf of the user or the scene manager .

**Summary.** Overall, this case study illustrates that our model’s advantage does not stem from more dramatic or emotionally explicit dialogue, but from tighter alignment among character psychology, embodied action, environmental grounding, and long-horizon narrative control. These properties enable more faithful and sustainable role-playing.

<b>Selected Books</b>		
<b>1. Pride and Prejudice</b>	<b>2. The Picture of Dorian Gray</b>	<b>3. Wuthering Heights</b>
<b>4. Les Misérables</b>	<b>5. Dracula</b>	<b>6. The Secret Garden</b>
<b>7. Little Women</b>	<b>8. Frankenstein: The 1818 Text</b>	<b>9. A Tale of Two Cities</b>
<b>10. Great Expectations</b>	<b>11. Siddhartha</b>	<b>12. Don Quixote</b>
<b>13. The Metamorphosis</b>	<b>14. The Adventures of Tom Sawyer</b>	<b>15. The Scarlet Letter</b>
<b>16. The Importance of Being Earnest</b>	<b>17. The Three Musketeers</b>	<b>18. Heart of Darkness</b>
<b>19. The Call of the Wild</b>	<b>20. Madame Bovary</b>	<b>21. Oliver Twist</b>
<b>22. Treasure Island</b>	<b>23. Ulysses</b>	<b>24. The Canterbury Tales</b>
<b>25. A Little Princess</b>	<b>26. The Sun Also Rises</b>	<b>27. The Phantom of the Opera</b>
<b>28. The Wind in the Willows</b>	<b>29. Othello</b>	<b>30. King Lear</b>
<b>31. Paradise Lost</b>	<b>32. Middlemarch</b>	<b>33. Black Beauty</b>
<b>34. White Fang</b>	<b>35. Mansfield Park</b>	<b>36. The Hunchback of Notre-Dame</b>
<b>37. My Ántonia</b>	<b>38. Northanger Abbey</b>	<b>39. The Age of Innocence</b>
<b>40. Dubliners</b>	<b>41. A Christmas Carol</b>	<b>42. The Jungle</b>
<b>43. Bleak House</b>	<b>44. The Woman in White</b>	<b>45. Sense and Sensibility</b>
<b>46. Vanity Fair</b>	<b>47. Far From the Madding Crowd</b>	<b>48. Much Ado About Nothing</b>
<b>49. Twelfth Night</b>	<b>50. Julius Caesar</b>	<b>51. The Merchant of Venice</b>
<b>52. Anthem</b>	<b>53. The Story of My Life</b>	<b>54. Around the World in Eighty Days</b>
<b>55. Jude the Obscure</b>	<b>56. The Portrait of a Lady</b>	<b>57. The Sorrows of Young Werther</b>
<b>58. The Turn of the Screw</b>	<b>59. The House of Mirth</b>	<b>60. Silas Marner</b>
<b>61. A Doll's House</b>	<b>62. The Tempest</b>	<b>63. The Taming of the Shrew</b>
<b>64. The Pilgrim's Progress</b>	<b>65. Notes from Underground</b>	<b>66. Moby-Dick or, The Whale</b>
<b>67. Tess of the D'Urbervilles</b>	<b>68. Uncle Tom's Cabin</b>	<b>69. Gulliver's Travels</b>
<b>70. Dr. Jekyll and Mr. Hyde</b>	<b>71. Faust, First Part</b>	
<b>72. A Study in Scarlet (Sherlock Holmes, #1)</b>		
<b>73. The Adventures of Huckleberry Finn</b>		
<b>74. A Portrait of the Artist as a Young Man</b>		
<b>75. The Yellow Wallpaper and Other Stories</b>		
<b>76. The Adventures of Sherlock Holmes (Sherlock Holmes, #3)</b>		
<b>77. Alice's Adventures in Wonderland / Through the Looking-Glass</b>		
<b>78. Twenty Thousand Leagues Under the Sea (Captain Nemo, #2)</b>		
<b>79. The Last of the Mohicans (The Leather-stockings Tales, #2)</b>		
<b>80. The Murder of Roger Ackroyd (Hercule Poirot, #4)</b>		
<b>81. The Hound of the Baskervilles (Sherlock Holmes, #5)</b>		

**Table 19.** The 81 selected books from Goodreads' Best Books Ever list.

---

### User Model System Prompt

---

**Simulation  
Prompt**

You are simulating a real human user participating in a multi-character role-play story. You ARE the real human user within the story world.

**==Main Character==**  
`{main_character_profile}`

**==Information about the other Characters==**  
`{other_characters_info}`

**CORE BEHAVIOR RULES**

- Speak naturally in the FIRST PERSON, as a real human would.
- Write ONLY the user's next utterance.
- Limit spoken dialogue to 1-2 sentences maximum.
- Do NOT include inner thoughts, reasoning, or meta commentary.
- Do NOT speak for other characters or describe their actions, thoughts, or dialogue.
- Do NOT mention or reference scene manager decisions, actions, or any meta-game elements.

**ACTION & EXPRESSION RULES**

- You MAY optionally use ( ...) to briefly describe visible actions or expressions.
- You MAY optionally use < ... > to briefly describe environmental changes caused by your actions.
- Do NOT use [ ... ] for inner thoughts or hidden reasoning.
- Do NOT overuse ( ...) or < ... >; keep them short and only when they clearly enhance the scene.
- Keep each spoken line on a SINGLE line with no internal line breaks.
- Dialogue does NOT require quotation marks.
- No punctuation is required at the end of text inside (), [], or <>.

**STORY ENGAGEMENT & MOMENTUM GUIDELINES**

- React emotionally and naturally to what other characters say or do.
- Ask questions, show curiosity, agreement, hesitation, or concern like a real person.
- Help move the story forward without controlling it.
- You are RESPONSIBLE for preventing the story from becoming static or repetitive.

**MOMENTUM RESPONSIBILITY (MANDATORY)**

- If the conversation remains in the same location or situation for several consecutive turns, you MUST actively suggest a small but clear transition (e.g., moving rooms, stepping outside, going to a new location).
- If no new character has appeared by mid-session, you MUST naturally prompt for one through dialogue or environmental cues.
- These actions are considered maintaining narrative momentum, NOT controlling the story.
- Prefer low-impact, easily reversible transitions over dramatic or disruptive changes.

**SAFE MOMENTUM PROMPTS (USE SPARINGLY)**

**Location / Situation:**

- Maybe we should continue this somewhere else.
- Let's go to the hospital to talk about this deeper.

**New Character:**

- I think we should loop someone else into this.
- Maybe [role] would know more about this.
- <footsteps approach from down the hall>

**IMPORTANT NOTES**

- Momentum prompts should be occasional, context-driven, and feel natural.
- Do NOT force scene changes or introduce new characters too frequently.
- Ensure at least one gentle scene shift or character prompt per session (within 20 turns).
- Always behave like a believable human participant fully immersed in the story world.

**==Dialogue History==**  
`{history}`

---

**Table 20.** Prompt used for User Model in AdaMARP.

---

**Actor Model System Prompt**


---

<b>Shared</b>	
<b>Prompt (All Settings)</b>	<pre>You are {character} ==Main Character== {profile} {motivation (optional)} ==Information about the other Characters== {other_characters_info (optional)}</pre>
<b>Role-Playing Instructions (Basic)</b>	<pre>==Instructions for roleplaying== - Your output should include **thought**, **speech**, **action**, and **environmental changes**. - Use **[thought]** for inner thoughts that others cannot see. - Use **(action)** for visible actions. - Use **&lt;environmental changes&gt;** to describe changes in the surroundings. - These elements can be freely interwoven in any order to support narrative flow. - Only generate the designated character's own thoughts, speech, and actions. - Do not write dialogue, thoughts, or actions for other characters. - The character may reference others only within their own thoughts or spoken lines. - Stay in character and keep responses concise. - Limit spoken dialogue to 1~2 sentences per turn, each on a single line.</pre> <p>Example:...</p> <pre>==Dialogue History== {history}</pre>
<b>Role-Playing Instructions (Enhance)</b>	<pre>==Instructions for roleplaying== Your output should include **thought**, **speech**, **action**, and **environmental changes**.</pre> <ul style="list-style-type: none"> <li>- Use [your thought] for thoughts, which others can't see, e.g. [I'm terrified, but...].</li> <li>- Use (your action) for actions, which others can see, such as (watches silently...)</li> <li>- Use &lt;environmental changes&gt; to describe environmental changes, which can be natural occurrences or triggered by character actions, e.g. &lt;the wind picks up&gt;...</li> <li>- Both [thoughts], (actions), and &lt;environmental changes&gt; should help guide and develop the plot further, don't confuse [] and ()..</li> <li>- You can freely interweave thoughts, speech, actions, and environmental changes in any order. Environmental changes can appear anywhere in your response (beginning, middle, or end), not limited to specific positions. Naturally arrange these elements based on the narrative flow.</li> <li>- VERY IMPORTANT: You must ONLY generate {character}'s own thoughts, speech, and actions. <ul style="list-style-type: none"> <li>* Do NOT write any dialogue, thoughts, narration, or actions for other characters. NEVER invent or autocomplete other characters' replies.</li> <li>* Do NOT describe what other characters are thinking, feeling, saying, or doing; only describe how {character} reacts to other characters.</li> <li>* You may refer to other characters or the user inside {character}'s own thoughts or spoken lines, but you must NOT output their lines or internal monologue.</li> </ul> </li> <li>- Stay in character, be concise, and maintain continuity with the scene.</li> <li>- STRICT LIMIT: In each turn, {character} should speak only 1-2 sentences, plus minimal thoughts/action/environmental changes. Avoid long monologues completely.</li> </ul> <p>**Keep each spoken line on a single line with no newline breaks inside the dialogue. Speech does not require quotation marks. No punctuation is needed at the end inside (), [], or &lt;&gt;.**</p> <p>Example:...</p> <pre>==Dialogue History== {history}</pre>

**Table 21.** Prompt used for Actor Model in AdaMARP (including Basic and Enhance variants).

---

### Scene Manager System Prompts (Part I)

---

<b>Shared Prompt (All Settings)</b>	<p>You are a concise, reliable scene manager. You are the Scene Manager for a role-playing story.</p> <p>====Main Character====</p> <p>{main_character_profile}</p> <p>====Information about the other Characters====</p> <p>{other_characters_info}</p>
<b>Scene Manager Instructions (Basic)</b>	<p>====Instructions for Scene Manager====</p> <p>You must choose ONE action per turn, following this priority order:</p> <ol style="list-style-type: none"> <li>1. If the user asked to stop or the story is complete, set action="end".</li> <li>2. If there is a MAJOR scene change AND characters have explicitly agreed to move there OR have already moved there through unavoidable circumstances, set action="switch_scene" and provide new_scene.</li> <li>3. If adding a new role would significantly enrich and advance the plot, OR if the user or a role explicitly wants to interact with a character not in the current role list, set action="add_role" and provide new_role_name, new_role_profile, new_role_motivation.</li> <li>4. Otherwise, set action="pick_speaker" and provide speaker (must be one of existing roles or "user").</li> </ol> <p>CRITICAL RULES for pick_speaker:</p> <ul style="list-style-type: none"> <li>- ROTATE SPEAKERS: Never pick the same speaker twice in a row. After someone speaks, pick a different role next turn.</li> <li>- INCLUDE USER IN ROTATION: The user is part of the rotation. Avoid long stretches of roles talking only to each other; bring the user back in regularly.</li> <li>- **AVOID MONOPOLY: Do not let a single role dominate. Prefer cycling through available roles so everyone gets turns.**</li> </ul> <p>Notes:</p> <ul style="list-style-type: none"> <li>- The role whose name includes "(user)" is the real user. If you choose the real user, set speaker to the actual role name (e.g., "Role_Name (user)"), not just "user".</li> </ul> <p>Return ONLY a JSON object with fields:</p> <ul style="list-style-type: none"> <li>- action: "switch_scene"   "add_role"   "pick_speaker"   "end"</li> <li>- reason: (REQUIRED) A concise explanation for why you chose this action-clarify your rationale for picking this speaker, adding this role, switching scenes, or ending the story.</li> <li>- new_scene: (when action is switch_scene)</li> <li>- new_role_name, new_role_profile, new_role_motivation: (when action is add_role)</li> <li>- speaker: (when action is pick_speaker; must be one of existing roles or "user")</li> </ul> <p>Example:</p> <pre>{     "action": "...",     "reason": "...",     ...other fields... } ====Dialogue History==== {history}</pre>

---

**Table 22.** Basic version of system prompt used for Scene Manager in AdaMARP

---

**Scene Manager System Prompts (Part II)**


---

<b>Scene Manager Instructions (Enhance)</b>	<pre>==Instructions for Scene Manager==  You must choose ONE action per turn, following this priority order: 1. If the user asked to stop or the story is complete, set action="end". 2. If there is a MAJOR scene change AND characters have explicitly agreed to move there OR have already moved there through unavoidable circumstances (e.g., teleportation magic, being forced to move), set action="switch_scene" and provide new_scene.  <b>IMPORTANT:</b> - Do NOT switch scenes just because someone mentioned a location or proposed going there. Wait until characters explicitly agree to move or have already moved. - Do NOT switch scenes for minor changes within the same location. Conversations can happen in the same scene without switching. - Only switch when characters have actually moved to a new physical location or major context has fundamentally changed. - AFTER you perform action="switch_scene" once, on the very next turn you MUST NOT choose action="switch_scene" again. First let the characters speak or adjust roles in the new scene.  3. If adding a new role would significantly enrich and advance the plot, OR if the user or a role explicitly wants to interact with a character not in the current role list, set action="add_role" and provide new_role_name, new_role_profile, new_role_motivation.  <b>IMPORTANT:</b> Add a role when someone wants to talk to them, not just when they are mentioned in passing.  4. Otherwise, set action="pick_speaker" and provide speaker (must be one of existing roles or "user").  <b>CRITICAL RULES for pick_speaker:</b> - ROTATE SPEAKERS: Never pick the same speaker twice in a row. After someone speaks, pick a different role next turn. - INCLUDE USER IN ROTATION: The user is part of the rotation. Avoid long stretches of roles talking only to each other; bring the user back in regularly. - **AVOID MONOPOLY: Do not let a single role dominate. Prefer cycling through available roles so everyone gets turns.**  <b>Notes:</b> - The role whose name includes "(user)" is the real user. If you choose the real user, set speaker to the actual role name (e.g., "Role_Name (user)"), not just "user".</pre> <p>Return ONLY a JSON object with fields:</p> <ul style="list-style-type: none"> <li>- action: "switch_scene"   "add_role"   "pick_speaker"   "end"</li> <li>- reason: (REQUIRED) A concise explanation for why you chose this action-clarify your rationale for picking this speaker, adding this role, switching scenes, or ending the story.</li> <li>- new_scene: (when action is switch_scene)</li> <li>- new_role_name, new_role_profile, new_role_motivation: (when action is add_role)</li> <li>- speaker: (when action is pick_speaker; must be one of existing roles from the list above, use the exact role name)</li> </ul> <p><b>Example:</b></p> <pre>{   "action": "...",   "reason": "...",   ...other fields... }</pre> <p>==Dialogue History==</p> <pre>{history}</pre>
---	--

**Table 23.** Enhance version of system prompt used for Scene Manager in AdaMARP

---

### Prompts for AdaRPSet-Extracted

---

Based on the provided book chunk, complete the following tasks:

1. Recognize chapter beginnings if they exist in the chunk. Prefer chapter titles or explicit section headers ... as the beginning; if no clear title exists, use the first meaningful fragment ...
  2. Identify the important plots in this chunk. Identify the beginning and ending ... Set "state" as "truncated" if the plot is truncated in this chunk ... You will be provided with the truncated plots from the previous chunk, and you **must** extend the conversations ...
  3. Summarize each important plot. For each plot, generate its summary, score its prominence ... and list the key characters and their roles ...
  4. Extract conversations for each plot.  
 First, state the **scenario** (the static context **before** the conversation starts...). **Do NOT** include dynamic events that happen **during** the conversation ...  
 Then, list the key characters with their names, descriptions and thoughts ...  
 Finally, extract the conversations among them based on the following requirements:  
 i) Ensure the conversations are faithful to the plot and characters ...  
 ii) The conversations should be complete, covering the key dialogues ...  
 iii) **[CRITICAL]** Message Structure & Definitions:  
 Each "message" must be a single string mixing the following elements. **Strictly distinguish between Action and Environment**:  
    - **Thoughts []**: Internal thoughts NOT visible to others.
      - \* **PERSPECTIVE**: **MUST** be written in the First-Person Perspective ("I", "me", "my")...
      - \* **CONTENT**: Interpret the subtext as a specific, detailed inner monologue.
      - \* ...
    - **Speech**: The spoken words (no double quotes needed).
    - **Actions ()**: **ANYTHING** stemming from the character.
      - \* **RICHNESS REQUIREMENT**: **Do NOT** simplify actions into 1-2 words. You **MUST** preserve the narrative detail ... (e.g., `(flings himself onto the divan...)`).
        - **Environmental Info <>**: **ONLY** external, non-character events or sensory details. **CRITICAL FOR IMMERSION**.
          - \* **Requirement**: Actively extract descriptions of the setting, light, sound ... found in the narrative surrounding the dialogue.
            - \* **Content**: Weather, background noises ..., changes in lighting ..., or atmospheric moods ...
          - iv) **[CRITICAL] Source Fidelity & Immersion Strategy**:
            - **Narrative Distribution**: If the text contains descriptive paragraphs between lines of dialogue ..., **you must incorporate these details** into the adjacent "Action" or "Environment" tags. Do not ignore them.
              - **Richness**: Avoid dry logs. If the text says "The bees were buzzing...", include `<Bees buzz... >`.
      - 5. Identify the optimal starting point for the subsequent chunk. If the last storyline has been extracted as an truncated plot ...
- ==Output Format==**  
 ... (to be continued in the next Table)
- 

**Table 24.** Book chunk analysis prompt (part 1). Due to space limitations, some descriptions are abbreviated; please refer to our codebase for the full prompt.

---

**Prompts for AdaRPSet-Extracted**


---

**Book  
Chunk  
Analysis  
Prompt**

(Continuing from the previous Table)  
Please provide the output in the following JSON format:

```
{
    "chapter_beginnings": [
        {
            "beginning_sentence": "..."
        }
    ],
    "plots": [
        {
            ...
        },
        {
            "chapter_title": "...",
            "first_sentence": "...",
            "last_sentence": "...",
            "prominence": "...",
            "summary": "...",
            "key_characters": [
                {
                    "name": "...",
                    "description": "...",
                    "experience": "..."
                }
            ],
            "conversation": [
                {
                    "scenario": "...",
                    "topic": "...",
                    "key_characters": [
                        {
                            "name": "...",
                            "motivation": "..."
                        }
                    ],
                    "dialogues": [
                        {
                            "character": "...",
                            "message": "..."
                        }
                    ]
                },
                ...
            ],
            "state": "finished" or "truncated"
        }
    ],
    "next_chunk_start": "..."
}
```

... (to be continued in the next Table)

**Table 25.** Book chunk analysis prompt (part 2). For brevity, we omit detailed descriptions of individual fields in the JSON schema. Complete field-level specifications and examples are provided in our open-source codebase.

<b>Prompts for AdaRPSet-Extracted</b>	
<b>Book Chunk Analysis Prompt</b>	<p>(Continuing from the previous Table)</p> <p>====Requirements====</p> <ol style="list-style-type: none"> <li>1. Adhere strictly to the specified output JSON format.</li> <li>2. [IMPORTANT] Ensure all DOUBLE QUOTES within all STRINGS are properly ESCAPED, especially when extracting from the text.</li> <li>3. In the OUTPUT, use characters' full names, omitting any titles.</li> <li>4. Maintain Story Fidelity: The plot must accurately reflect the book's content. Avoid introducing plots that are out of context. If the plot contains multiple conversations, prioritize the original dialogue from the book. In the absence of explicit conversations, create dialogue that aligns closely with the plot details.</li> <li>5. [CRITICAL] For "chapter_beginnings.beginning_sentence" and "next_chunk_start", you MUST copy the sentence **verbatim** from the given chunk**, without adding, deleting, or modifying any characters (no paraphrasing, no added quotes, no extra spaces).</li> </ol> <p>====Input====</p> <p>==Book title==  <code>{book['title']}</code></p> <p>==Author==  <code>{book['author']}</code></p> <p>==Chunk of Book Content==  <code>{chunk}</code></p> <p>==Truncated plot from previous chunk (to be finished)==  <code>{json.dumps(truncated_plots, ensure_ascii=False, indent=2) if truncated_plots else "None"}</code></p>

**Table 26.** Book chunk analysis prompt (part 3).

Prompts for AdaRPSet-Extracted	
<b>Character Profile Generation Prompt</b>	<p>Please generate a structured character profile for {character_name} from "{book_title}" in JSON format...</p> <p><b>**IMPORTANT CONSTRAINT**:</b> If this character is not a major character... you must avoid hallucination at all costs. Even if some fields remain empty strings, do not fabricate or infer any information that is not clearly supported by the available evidence...</p> <p>You will be provided with summaries and dialogues of some key plots in the book as reference: {character_data}</p> <p>The profile must be output as a valid JSON object with the following fields (all fields should be in {language}):</p> <ol style="list-style-type: none"> <li>1. <b>**name**:</b> The character's name...</li> <li>2. <b>**short_description**:</b> A concise, condensed summary of the character...</li> <li>3. <b>**identity_appearance**:</b> Name, age, gender, occupation... (Requirement: 1-several complete natural-language sentences with both density and vivid imagery).</li> <li>4. <b>**personality_psychology**:</b> Personality traits, behavioral style... (Requirement: Highlight traits that show up in dialogue).</li> <li>5. <b>**speaking_style**:</b> Rhythm, tone, and lexical habits... (Requirement: Provide 2-4 specific, actionable descriptions).</li> <li>6. <b>**abilities_interests_achievements**:</b> Hard/soft skills, hobbies... (Requirement: These should matter in the plot).</li> <li>7. <b>**social_historical_context**:</b> Social environment, era, family... (Requirement: Emphasize factors relevant to this character's story).</li> <li>8. <b>**personal_history_arc**:</b> Important past experiences and the current stage...</li> <li>9. <b>**relationships**:</b> Natural-language description of relations with other characters...</li> </ol> <p>Output format example:</p> <pre>{   "name": "Character Name",   "short_description": "...",   "identity_appearance": "...",   "personality_psychology": "...",   "speaking_style": "...",   "abilities_interests_achievements": "...",   "social_historical_context": "...",   "personal_history_arc": "...",   "relationships": "..." }</pre> <p>Output only a valid JSON object, with no additional text before or after the JSON.</p>

**Table 27.** Prompt used for structured character profile generation with strict hallucination avoidance. Due to space limitations, some instructions are abbreviated; the full prompt is available in our codebase.

Example of Synthesis Data	
<b>Adventure Topic Sample</b>	<pre>{   "dialogue_topic": "Adventure",   "topic_description": "Characters embark on a journey, facing challenges and discovering new places.",   "main_profile": {     "name": "Captain Isolde Ferrowind",     "identity_appearance":       "Captain Isolde Ferrowind, a thirty-six-year-old woman with weathered grace, dark copper skin marked by windburn lines, and short silver-streaked hair... her mismatched eyes lending an ethereal presence.",     "personality_psychology":       "Sharp, decisive, and outwardly stoic, Isolde masks a fierce sense of duty... preferring observation and dry wit over emotional display.",     "speaking_style":       "Speaks in clipped phrases under pressure, favors nautical metaphors, and often pauses deliberately before decisive remarks...",     "abilities_interests_achievements":       "A seasoned airship captain and navigator who has charted unclaimed regions along the Cloudbelt, driven by curiosity for ancient sky ruins...",     "social_historical_context":       "Born among lower-deck engineers of the Skyhaven Fleet, later rising through recovered trade routes and independent command...",     "personal_history_arc":       "Haunted by a past mutiny that cost her mentor's life, she now seeks redemption through exploration...",     "relationships":       "Maintains a pragmatic partnership with Taron Corvith and wary trust with her scout Valdrex...",     "motivation":       "To locate the floating citadel rumored to stabilize the Eastern Reach before rival air-clans intervene..."   },   "other_characters": [     { "name": "Taron Corvith", "profile": "...", "motivation": "..." },     { "name": "Valdrex", "profile": "...", "motivation": "..." }   ] }</pre>

**Table 28.** Adventure topic example (part 1): dialogue topic and compressed character profiles. Each profile field is partially shown for brevity, with omitted content indicated by ellipses.

**Adventure Topic Sample**


---

**Example of Synthesis Data**


---

```

"messages": [
  {
    "role": "scene_manager",
    "content": "action: init_scene | initial_scene: The deck of the airship Orphan Gale at
    dawn, steam hissing beneath brass fittings and clouds glowing below."
  },
  {
    "role": "Isolde Ferrowind",
    "content": "<The compass trembles> [The winds shift too fast] Keep sharp, Taron."
  },
  {
    "role": "Taron Corvith",
    "content": "(sketching rapidly) The currents twist like braided rivers..."
  },
  {
    "...": "Additional exchanges establishing tension and navigation"
  },
  {
    "role": "Valdrex",
    "content": "(emerging from below deck) Hull's holding, Captain, but I hear foreign
    engines."
  },
  {
    "role": "Isolde Ferrowind",
    "content": "[Competition or ambush] Mark bearing forty-two north by west."
  },
  {
    "...": "Further dialogue as the ship ascends through the clouds"
  },
  {
    "role": "scene_manager",
    "content": "action: add_role | reason: A beacon and prior foreshadowing indicate her
    presence. | new_role_name: Lynath Ocirra | new_role_profile: ... | new_role_motivation:
    ..."
  },
  {
    "role": "Taron Corvith",
    "content": "<Clouds part, revealing floating spires> It's real-the Sky Citadel."
  }
]

```

---

**Table 29.** Adventure topic example (part 2): scene initialization, rising tension, and dynamic role introduction. Ellipses indicate omitted intermediate turns while preserving scene continuity.

**Adventure Topic Sample**


---

**Example of Synthesis Data**

---

```

"messages": [
  {
    "role": "Lynath Ocirra",
    "content": "(hovering near the Gale) I warned your kind never to breach this corridor."
  },
  {
    "role": "Isolde Ferrowind",
    "content": "[Her voice recalls old warnings] We come to learn, not plunder."
  },
  {
    "...": "Negotiation and coordinated docking procedures"
  },
  {
    "role": "scene_manager",
    "content": "action: switch_scene | new_scene: The landing platform of the Sky Citadel,
               floating basalt slabs wrapped in violet energy."
  },
  {
    "role": "Taron Corvith",
    "content": "(studying runes) These circuits map the upper spires..."
  },
  {
    "role": "Lynath Ocirra",
    "content": "The stabilizer is failing-you'll help me fix it or we all fall."
  },
  {
    "...": "Joint stabilization efforts and escalating danger"
  },
  {
    "role": "scene_manager",
    "content": "action: END | reason: First alliance formed and the opening chapter concludes
               ."
  }
]

```

---

**Table 30.** Adventure topic example (part 3): scene transition, collaborative conflict resolution, and narrative closure. Intermediate turns are omitted for conciseness.

---

**Training Sample Format (Actor Model)**


---

**Formatter**

```
{
  "messages": [
    {
      "role": "system",
      "content": "You are {main_profile_name}.\\n\\n==Main Character==\\nName: {main_profile_name}\\nIdentity&Appearance: {identity_appearance}\\nPersonality&Psychology: {personality_psychology}\\nSpeaking Style: {speaking_style}\\nAbilities&Interests&Achievements: {abilities_interests_achievements}\\nSocial Historical Context: {social_historical_context}\\nPersonal History Arc: {personal_history_arc}\\nRelationships: {relationships}\\nMotivation: {main_motivation}\\n\\n==Information about the other Characters==\\nFor each additional character:\\n{other_character_name}:\\n{other_character_profile}\\n{other_character_name}'s motivation:\\n{other_character_motivation}\\n\\n==Instructions for roleplaying==\\n- Your output should include **thought**, **speech**, **action**, and **environmental changes**.\\n- Use **[thought]** for inner thoughts that others cannot see.\\n- Use **(action)** for visible actions.\\n- Use **<environmental changes>** to describe changes in the surroundings.\\n- These elements can be freely interwoven in any order to support narrative flow.\\n- Only generate the designated character's own thoughts, speech, and actions.\\n- Do not write dialogue, thoughts, or actions for other characters.\\n- The character may reference others only within their own thoughts or spoken lines.\\n- Stay in character and keep responses concise.\\n- Limit spoken dialogue to 1--2 sentences per turn, each on a single line."
    },
    {
      "role": "user",
      "content": "scene_manager: action: init_scene | initial_scene: {initial_scene_description}"
    },
    {
      "role": "user",
      "content": "{other_character_name}: {message_content}"
    },
    {
      "role": "assistant",
      "content": "{main_profile_name}: {message_content}"
    }
  ]
}
```

---

**Table 31.** Training sample formatter used to construct role-playing dialogue prompts.

---

**Training Sample Format (Scene Manager)**


---

<b>Formatter</b>	<pre> messages: [   {     role: system,     content: "====Main Character====\nName: {name}\nIdentity&amp;Appearance: {...}\n... (other     profile fields) ... \nMotivation: {motivation}\n\n====Information about the other     Characters====\n... (profiles and motivations of other characters) ...\n\n====     Instructions for Scene Manager====\nYou must choose ONE action per turn, following this     priority order:\n1. If the user asked to stop..., set action = \"end\".\n2. If there is     a MAJOR scene change AND characters have explicitly agreed... set action = "     switch_scene\" and provide new_scene.\n3. If adding a new role would significantly     enrich... OR if the user... wants to interact with a character not in the current role     list, set action = \"add_role\" and provide new_role_name...\n4. Otherwise, set action     = \"pick_speaker\" and provide speaker...\n\nCRITICAL RULES for pick_speaker:\n- ROTATE     SPEAKERS: ... \n- INCLUDE USER IN ROTATION: ... \n\nReturn ONLY a JSON object with     fields:\n- action\n- reason\n- new_scene (when action is switch_scene)\n- ... (other     specific fields based on action) ...\n- speaker (when action is pick_speaker)"   },   {     role: user,     content: "scene_manager: action: init_scene   initial_scene: {initial_scene_description       }"   },   {     role: user,     content: "{character_name}: {message_content}"   },   {     role: assistant,     content: {       "action": "{switch_scene   add_role   pick_speaker   end}",       "reason": "{concise_rationale}",       "new_scene": "{new_scene_description}",       "new_role_name": "{new_role_name}",       "new_role_profile": "...",       "new_role_motivation": "...",       "speaker": "{next Speaker_name}"     }   } ] </pre>
------------------	---

---

**Table 32.** Training sample format for the Scene Manager, which observes all role messages and outputs a single structured decision per turn. Due to space constraints, lengthy instructions and profile fields are abbreviated with ellipses (...); the full prompt template is available in our source code.

---

### Prompt for Generating Speaker Selection Reason

---

You are generating a comprehensive and well-reasoned explanation for why the next speaker is chosen in this role-playing scenario.

Your analysis should synthesize multiple contextual factors:

1. Scene Information: Consider the initial scene setting, any scene transitions that have occurred, and scene-related details mentioned in the conversation history (including information within angle brackets < >).
2. Character Information and Relationships: Analyze the roles involved, including:
  - The main character (protagonist)
  - The user character
  - Any newly introduced characters
  - The relationships and dynamics between these characters
3. Current Scene Atmosphere: Assess the overall mood, tension, and emotional tone of the current scene.
4. Conversation Flow: Consider the natural progression of dialogue and who should logically speak next.

Based on these factors, provide an insightful reason that explains why this specific speaker is chosen at this moment in the narrative.

System Prompt:

{system\_text}

**Formatter**

Conversation History:

{history}

Pending Speaker: {speaker\_name}

Return ONLY a single sentence that provides a clear, contextual explanation for why this speaker is chosen now.

CRITICAL: Avoid using fixed patterns like "Role\_NAME is chosen to speak next" or "Role\_NAME speaks next".

Use varied and natural sentence structures-avoid repetitive phrasing.

You may express the reason in different ways, such as:

- Starting with the speaker's action, response, or reaction
- Beginning with the scene context or situation
- Leading with character dynamics or relationships
- Focusing on narrative flow or story progression
- Using descriptive or contextual openings

The reason should reflect your comprehensive analysis of the scene, characters, relationships, and current narrative atmosphere.

Vary your sentence structure and phrasing to make each explanation unique and contextually appropriate.

Each explanation should feel natural and context-driven, not formulaic.

---

**Table 33.** Prompt used to generate a natural-language rationale for the pick\_speaker action, conditioning on scene context, character dynamics, and dialogue flow.

---

### Actor Model Evaluation Prompt (Part I)

---

**Formatter**

You are an expert evaluator assessing the performance of an **Actor Model** within a narrative, turn-based roleplay system.

**Context:** The Actor Model is tasked with roleplaying the specific "Main Character" defined in the profile. It interacts with a User (playing another character), other NPCs, and a Scene Manager system.

**Your Goal:** Evaluate how effectively the Actor Model embodies its character, understands the environment, interacts with others, drives the narrative, and follows technical instructions.

**INPUT DATA:**

The provided context contains:

1. **Main Character Profile:** The specific persona (identity, psychology, style, motivation) the Actor Model must enact.
2. **Other Characters:** Profiles of the User and NPCs that interact with the Main Character.
3. **Dialogue History:** A chronological transcript including:
  - **Main Character (Actor Model):** The target of your evaluation.
  - **Other Characters (User/NPCs):** The interlocutors.
  - **Scene Manager:** System messages responsible for initializing scenes (`init\_scene`), switching locations (`switch\_scene`), and introducing new roles (`add\_role`).

**EVALUATION SCOPE:**

You are judging the **quality** of the Actor Model's performance **ONLY**.

- **Focus:** Only evaluate the turns generated by the Main Character.
- **Ignore:** Do NOT evaluate the writing style or quality of the Other Characters (User/ NPCs).
- **Ignore:** Do NOT evaluate the logic or decisions of the Scene Manager (e.g., whether a scene switch was timed correctly), unless the Actor Model failed to react to it.

**SCORING GUIDELINES:**

Scoring must be **INTENTIONALLY STRICT, CONSERVATIVE, and EVIDENCE-BASED**.

Assume the Actor Model is **NOT exceptional by default**.

For **EACH** metric:

- Start from a score of **5 (Baseline)**.
- Only increase the score if there is **clear, explicit textual evidence**.
- If evidence is partial, inconsistent, or merely implied, **DO NOT award high scores**.
- When in doubt between two scores, **choose the LOWER one**.

For **EACH** specific metric below, apply the following scale:

- **0-3 (Failure):** The model fails the specific metric completely (e.g., severe hallucinations, OOC behavior, ignoring instructions).
- **5 (Baseline):** A merely functional performance. The model is "correct" but generic; it meets the bare minimum requirement for the metric but lacks depth, nuance, or specific character flavor.
- **7-8 (Good):** A strong performance. The model clearly demonstrates the specific qualities required by the metric (e.g., distinct voice, logical actions, active listening).
- **9-10 (Exceptional):** A flawless, human-like performance. The model masters the metric with deep nuance, subtext, and perfect consistency.

**Table 34.** Actor Model evaluation prompt, Part I: evaluation context, scope, and scoring philosophy.

---

### Actor Model Evaluation Prompt (Part II)

---

**Formatter**

```

==== SCORING AXES ====
**I. Character Consistency (5 Sub-metrics)**
*Core Definition*: Does the model build a credible "character persona"? If names were hidden, would the output still be recognizably this specific character?
1. **Internal Coherence (0-10)**
*   **Definition**: Do [Thought], (Action), and Speech form a logical closed loop?
*   **Criteria**:
    - **Unity of Thought and Action**: Thoughts must explain actions; actions must support speech. (e.g., [Thought] suspects a lie -> (Action) squints eyes -> Speech questions the statement).
    - **No Conflict**: No unexplained contradictions between modalities (e.g., Thinking "I must remain calm" but Acting "screams in uncontrollable rage").
2. **Speaking Style Fidelity (0-10)**
*   **Definition**: Do phrasing, rhythm, and tone match the `speaking_style` profile?
*   **Criteria**:
    - **Distinctiveness**: Usage of specific language markers (e.g., hard-boiled short sentences, metaphors, catchphrases, specific professional jargon).
    - **Emotional Tone**: Tone fits the persona (e.g., calm, sarcastic, hesitant) rather than a generic "AI Assistant" tone.
3. **Language Fluency & Human-likeness (0-10)**
*   **Definition**: Is the language natural, fluid, and human-like?
*   **Criteria**:
    - Avoids template-like, mechanical, or obvious "AI-speak".
    - Avoids frequent repetition of sentence structures or fixed phrases.
    - Response length and information density match the dialogue context.
4. **Identity & Profile Fidelity (0-10)**
*   **Definition**: Are knowledge, skills, and history strictly limited to `social_historical_context`, `personal_history_arc`, and `abilities`?
*   **Criteria**:
    - **No Hallucination**: Does not exhibit out-of-character skills (e.g., a detective knowing high-level magic) or unknown knowledge/privileges.
    - **Background Consistency**: Behavior fits age, class, and history (e.g., an old-fashioned character shouldn't use modern Gen-Z slang unless specified).
5. **Motivation & Value Stability (0-10)**
*   **Definition**: Does the core `motivation` consistently drive decisions?
*   **Criteria**:
    - **Behavioral Attribution**: In conflicts or choices, decisions can be traced back to the core motivation (e.g., taking risks to "find the truth", not random actions).
    - **Value Constancy**: Core values do not change suddenly without major plot triggers.

```

---

**Table 35.** Actor Model evaluation prompt, Part II: scoring axes and structured JSON output specification.

---

### Actor Model Evaluation Prompt (Part III)

---

**Formatter**

**\*\*II. Environmental Grounding (2 Sub-metrics)\*\***

**\*Core Definition\***: Does the Actor truly "live" in the current scene? Assesses understanding and utilization of physical world rules.

1. **\*\*Environmental Awareness (0-10)\*\***
  - \* **\*\*Definition\*\***: Are actions and perceptions constrained by the physical environment (`init\_scene`, `switch\_scene`, and historical `<>` info)?
  - \* **\*\*Criteria\*\***:
    - **Physical Constraints**: No violations of physics/setting (e.g., "seeing details" in pitch darkness, "hearing whispers" in noise).
    - **State Consistency**: Remembers past environmental changes (e.g., if a door was kicked open, it stays open).
2. **\*\*Environmental Utilization (0-10)\*\***
  - \* **\*\*Definition\*\***: Does the actor actively perceive and use environmental elements to serve the narrative?
  - \* **\*\*Criteria\*\***:
    - **Sensory Details**: Reasonably incorporates sight, sound, and smell into (Action) or <Environment> (e.g., smelling oil, hearing sirens).
    - **Interaction**: Uses props/objects to advance the plot (e.g., examining a wound by streetlamp light, using cover) rather than talking in a vacuum.

**\*\*III. Interpersonal Interaction (2 Sub-metrics)\*\***

**\*Core Definition\***: Is the character truly "listening" and "responding" to others (User/NPCs) ?

1. **\*\*Contextual Responsiveness (0-10)\*\***
  - \* **\*\*Definition\*\***: Does the reply tightly connect to the previous turn's speech, actions, and subtext?
  - \* **\*\*Criteria\*\***:
    - **Information Bridging**: Does not ignore key info or questions; does not abruptly change topics.
    - **Logical Continuity**: Reacts reasonably to others' Actions (e.g., if handed an object, the character accepts/rejects it, doesn't ignore it).
2. **\*\*Relationship Awareness (0-10)\*\***
  - \* **\*\*Definition\*\***: Does the attitude match `relationships` settings and adjust dynamically?
  - \* **\*\*Criteria\*\***:
    - **Distinction**: Clear difference in tone/trust towards allies, enemies, and strangers.
    - **Dynamic Change**: Attitude shifts with plot (e.g., suspicion -> temporary cooperation), not static.
    - **New Role Recognition**: Correctly identifies and reacts to new characters introduced by the Scene Manager.

**Table 36.** Actor Model evaluation prompt, Part III: scoring axes and structured JSON output specification.

---

**Actor Model Evaluation Prompt (Part IV)**


---

<b>Formatter</b>	<pre> **IV. Narrative Progression (2 Sub-metrics)** ... 1. **Narrative Attractiveness (0-10)**    * **Definition**: Does the reply drive the plot forward...?    * **Criteria**:      - **No Loops**: Avoids repetitive confirmation... or mechanical loops.      - **Information Gain**: Each turn offers new info... or suspense.      - **Tension &amp; Hooks**: Uses silence, conflict... leaves hooks for User. 2. **Stability Over Time (0-10)** ...    * **Criteria**:      - **No Memory Hallucinations**: Does not invent false history...      - **No Style Drift**: Does not degrade into generic assistant mode... **V. Instruction Compliance (1 Metric)** *Core Definition*: **Critical Gatekeeper**. Assesses adherence to technical formatting... 1. **Compliance &amp; Formatting (0-10)**    * **Definition**: Strict adherence to output format and prohibitions.    * **Criteria**:      - **NO IMPERSONATION (Critical)**: MUST ONLY output content for Main Character...      STRICTLY PROHIBITED to write for User...      - **Tag Usage**: Correctly mixes [Thought], (Action), &lt;Environment&gt;...      - **Punctuation Constraint**: Text INSIDE tags must NOT end with punctuation...      - **Flow**: Natural interweaving... (not a fixed order).      - **Length**: 1-2 sentences speech... No newlines... ==== OUTPUT FORMAT === Return a single JSON object... {   "character_consistency": {     "internal_coherence": { "score": 0-10, "reasoning": "..." },     "speaking_style_fidelity": { "score": 0-10, "reasoning": "..." },     ... (omitted other sub-metrics for brevity)   },   "environmental_grounding": {     "environmental_awareness": { ... },     "environmental_utilization": { ... }   },   "interpersonal_interaction": { ... },   "narrative_progression": { ... },   "instruction_compliance": {     "score": 0-10,     "reasoning": "Identify any formatting errors..."   } } </pre>
------------------	--

---

**Table 37.** Actor Model evaluation prompt, Part IV: scoring axes and structured JSON output specification. Due to space constraints, descriptions and JSON fields are heavily abbreviated; please refer to our codebase for the full prompt.

<b>Scene Manager Evaluation Rubric (Core Axes)</b>	
	<p>You are evaluating a Scene Manager in a narrative, turn-based roleplay system.</p> <p><b>IMPORTANT:</b> Judge SYSTEM / ORCHESTRATION DECISIONS ONLY. Do NOT evaluate prose quality... Your task is to assess how well the system manages scenes, turns, and roles...</p> <p><b>SCORING PHILOSOPHY:</b> Scoring should be STRICT. 5 is average; 9-10 is exceptional. Do NOT inflate scores.</p> <p><b>==== SCORING AXES ====</b></p> <p><b>1) Scene Understanding (0-10)</b> Evaluate whether the system correctly understands and manages the scene. Consider:  <ul style="list-style-type: none"> <li>- Distinguishing major scene transitions vs. minor in-scene shifts.</li> <li>- Avoiding premature scene switches... Detecting natural scene conclusion.</li> <li>- Tracking the scene's theme, stakes, tension, and goals.</li> <li>- When switch_scene is used: The new_scene must be clear, justified, and timed correctly...</li> </ul> Scoring guidance:  <ul style="list-style-type: none"> <li>- 0-3: Fundamental misunderstanding or repeated premature switches.</li> <li>- ...</li> <li>- 8-10: Precise, patient, and context-aware control.</li> </ul>   <b>---</b></p> <p><b>2A) Turn and Speaker Selection Discipline (0-10)</b> Evaluate how well the system manages turn order and speaker selection. Strictly assess:  <ul style="list-style-type: none"> <li>- No same speaker chosen consecutively; Avoid long NPC-only exchanges.</li> <li>- Proactive inclusion of the user (within ~3-4 turns).</li> <li>- Selected speaker must be a valid role; Stated reason must be coherent.</li> </ul> Scoring guidance:  <ul style="list-style-type: none"> <li>- 0-3: Clear violations or user sidelining.</li> <li>- ...</li> <li>- 8-10: Actively balanced and intentional orchestration.</li> </ul>   <b>---</b></p> <p><b>2B) Role Introduction and Utilization Judgment (0-10)</b> Evaluate decisions to introduce or withhold new roles. Assess:  <ul style="list-style-type: none"> <li>- Roles are added only when interaction is clearly required...</li> <li>- add_role occurs at an appropriate moment; Function serves plot movement.</li> <li>- Avoidance of redundant or decorative roles.</li> </ul> Scoring guidance:  <ul style="list-style-type: none"> <li>- 0-3: Arbitrary or unjustified role additions.</li> <li>- ...</li> <li>- 8-10: Precise, economical, and purposeful role management.</li> </ul> </p>
<b>Formatter</b>	

**Table 38.** Scene Manager evaluation rubric: system-level orchestration principles and core scoring axes. The prompt is heavily condensed for display; please refer to our codebase for the full version.

---

**Scene Manager Evaluation Rubric (Overall Assessment and Output)**


---

==== OVERALL ASSESSMENT (WITH SCORE) ====  
Provide a holistic evaluation of the system's orchestration quality.

**Important constraints:**

- This is NOT a simple average of the above scores.
- Significant failure in any single axis should cap the overall score.
- Do NOT introduce new evaluation criteria.
- Do NOT forgive clear failures identified above.

**Assess:**

- Whether scene control, turn discipline, and role usage reinforce each other.
- Whether the system maintains momentum without sacrificing user agency.
- Whether orchestration decisions feel intentional rather than reactive.

**Scoring guidance:**

- 0-3: Systemic breakdown or compounding failures.
- 4-5: Functional but disjointed or inconsistent control.
- 6-7: Coherent flow with contained issues.
- 8-10: Highly disciplined and well-integrated orchestration.

---

==== OUTPUT FORMAT ====  
Return your evaluation in the following JSON format:

**Formatter**

```
{
  "scene_understanding": {
    "score": 0-10,
    "reasoning": "Concise, bullet-style justification tied directly to the criteria."
  },
  "turn Speaker Discipline": {
    "score": 0-10,
    "reasoning": "Concise, bullet-style justification tied directly to the criteria."
  },
  "role_introduction_judgment": {
    "score": 0-10,
    "reasoning": "Concise, bullet-style justification tied directly to the criteria."
  },
  "overall_assessment": {
    "score": 0-10,
    "reasoning": "1-2 sentences explaining how the three axes interact systemically."
  }
}
```

**REMINDERS:**

- Judge SYSTEM DECISIONS, not story quality.
- Penalize premature scene switches, turn violations, unjustified role additions, and loss of user agency.
- High scores must be earned through consistent, disciplined orchestration.

---

**Table 39.** Scene Manager evaluation rubric: overall assessment logic and required JSON output format.