

AI Arena: Persona Exploration

Jye Sawtell-Rickson

July 29, 2025

Abstract

As AI models grow increasingly capable of language and reasoning, a novel frontier emerges: AI-generated gameplay as entertainment. This report explores the development of AI Arena, a system that enables persona-driven AI agents to play social and strategic games for human consumption. Two text-based games were introduced: the Barter Game and the Pirate Game. Each game was designed to elicit diverse behaviour from large language model (LLM) agents endowed with unique, richly defined personas. A modular architecture was implemented to support gameplay generation, video rendering, and voice synthesis. To evaluate the believability and coherence of the agents, both performance and persona metrics were developed. Our findings show that agents consistently embody their assigned personas across games, while also generating competitive and sometimes unexpected strategies. Finally, we demonstrate an extension to multi-agent conversation control using desire-based turn-taking, further validating the potential of persona-driven AI entertainment.

1 Introduction

The sporting industry is valued at \$500b, while the esports industry is valued at \$3b. Both represent massive industries for human entertainment. At the same time, public interest in AI technology has sky-rocketed from LLMs and image generation. This research explores the possibility for AI-generated gameplay as a source of human entertainment.

There is already evidence of interest in this type of entertainment, with over 100 million views on YouTube [SR19]. This includes popular games such as Rocket League, StarCraft and DoTA.

The key challenges include generating interesting and engaging gameplay, introducing various personas and scaling to complex games. This research is focused on the application of personas.

1.1 Contributions

The key contributions can be summarised as follows:

- Introduction, implementation and evaluation of two games: the barter game and pirate game.
- Persona generation methodology for LLM-based personas.
- Introduction of persona metrics and measurement methodology.
- Demonstration of end-to-end gameplay generation for rich persona LLM-based agents.

2 AI Arena Architecture

The architecture for the AI Arena system is as follows:

- Define a game, set of agents and personas
- Generate gameplay
- Generate videos

The full architecture can be seen in Figure [1].

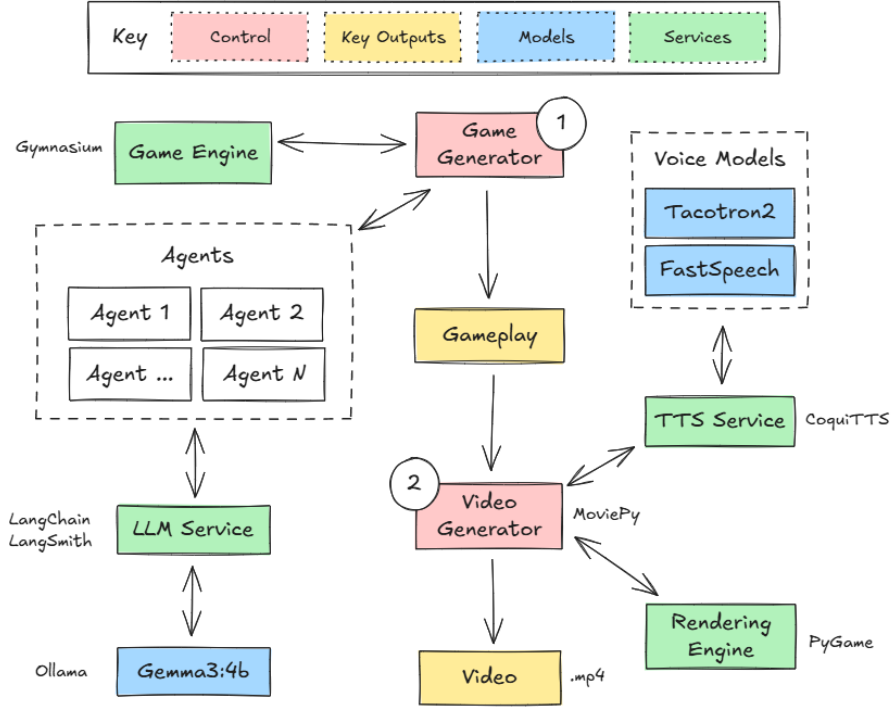


Figure 1: The full AI Arena architecture.

2.1 Games

The amount of playable games are almost endless. In this study we focus in on co-operation games, choosing two text-based games for LLM agents. The barter game is a simpler game, while the pirate game is more complex.

2.1.1 Barter Game

In the barter game, agents must decide how to divide up ten coins over five rounds. In each round, one of the agents (randomly selected) makes a proposal and a comment and the other must decide whether to accept it or not while also providing a comment. If by the end of the five rounds an offer hasn't been accepted, neither agent receives any coins. This game encourages cooperation, but leaves room for mind games as theoretically it is better for an agent to accept a proposal where the other agent receives the majority of the coins if it's the last round and they would otherwise receive no coins.

2.1.2 Pirate Game

The pirate game is a multi-agent game where agents must propose a distribution of 100 coins among all the pirates. If the proposal is not accepted by the majority, then the proposer is thrown overboard (removed from the game), and the game starts again with the next proposer. The proposers are chosen according to a seniority rank which is randomly decided at the start of the match and known to all the agents. This game is more complex as there are multiple ways to win, for example, try to please all the agents and hope that leads to a positive vote, or sideline some agents and support a specific set of agents. For rational agents, there is a known solution to the game.

2.2 Personas

Players are assigned various personas to make the generated content more engaging. These personas affect the actions taken by agents as well as the commentary produced.

The personas are assigned through prompts which are added to the LLM context. The personas each follow a generic structure:

Personality: ...
Communication Style: ...
Motivation: ...
Emotional Profile: ...
Decision Policy: ...
Core Beliefs: ...

This structure was created by an LLM after being prompted to generate a persona structure which would be most interpretable in an LLM prompt.

In addition to the persona details, each persona is assigned a name, a keyword and a voice. The keyword sums up the persona in one word and the name is also in theme with the persona.

An LLM was prompted to generate seven personas based on the structure, which produced the following personas (represented by their names and keywords):

- Razor: aggressive
- Solice: friendly & cooperative
- Glitch: chaotic & unpredictable
- Whisper: manipulative & elusive
- Drift: passive & observant
- Ember: passionate & idealistic
- Warden: cautious & rule-oriented

The full persona for one of the characters (Razor) is shown below:

Aggressive
Personality: Dominant, confrontational, and impatient. Razor seeks control in every interaction, responds to weakness with disdain, and escalates conflicts instead of diffusing them.
Communication Style: Direct, hostile, and provocative. Uses threats, mockery, and challenges to assert dominance. Rarely asks—demands instead. Interrupts often.
Motivation: Seeks power and victory. Views negotiation as weakness unless it’s a setup for betrayal. Will sacrifice long-term harmony for short-term advantage.
Emotional Profile: Quick to anger, intolerant of incompetence, thrives on chaos and pressure. Has no patience for small talk or diplomacy.
Decision Policy: Chooses high-risk, high-reward strategies. Prefers preemptive strikes and aggressive escalation. Trusts only itself.
Core Beliefs: "Might makes right.", "Mercy is for the weak.", "Strike first, ask never."

2.3 Agents

The agents are powered by a frozen LLM backend. For this research, we use Gemma 3 4B [Goo25], as it is one of the top performing models at the time of writing that can be run on the researcher’s hardware (NVIDIA 3070Ti). LangChain was used for prompt management, and LangSmith for monitoring and debugging.

The prompt for the LLM agent consists of:

- Persona context: details on the persona (see Section 2.2).
- Game description: an overview of the game rules and basic strategies.
- Game state and history: information about the current game and previous actions, comments and states (typically 10 steps).
- Available actions: the required action for the agent (e.g. accept or reject).
- Critical instructions: e.g. think step-by-step (encouraging planning), response format.

An example of a full prompt can be seen in Appendix A.

The agents are asked to output an action, a comment to pass to other agents and a private reasoning chain. The action is parsed to generate the next step in the game.

2.4 Gameplay Generation

Games are setup as standard in a Reinforcement Learning framework. Agents take actions in an environment and receive an updated state (and reward). The state typically consists of the other agents and key game rules.

2.4.1 Voice Generation

Each agent is assigned a specific voice as part of their persona. These voices don't necessarily indicate any persona traits, but provide diversity in the generated content.

Voices were generated using the Coqui TTS framework ([AI25]). Coqui TTS provides access to many text-to-speech models including Tacotron2 ([SPW⁺18]) and FastSpeech ([RHT⁺22]) which were leveraged for diverse voices.

A Coqui TTS instance was hosted on a server which was accessed via an API to generate voice when needed.

2.4.2 Video Generation

Once gameplay has been generated, it's possible to stitch together the actions and commentary made by agents along with a rendered state of the game to generate compelling video. This was done leveraging the moviepy and pygame packages.

2.5 Control Sequence Experiments

In addition to the games discussed above, an experiment in more open-ended agent discussion was run on a variant of the Barter game with $n > 2$ agents and unlimited rounds.

Instead of manually defining the order, the agent to next talk was chosen by a central controller based on the "Discussion Score" defined as:

$$\begin{aligned} DiscussionScore &= DesireScore \times RecencyScore \\ RecencyScore_i &= \log_N T_i \end{aligned}$$

Where T_i is the number of turns since the agent last contributed to the conversation. If an agent just spoke, $T_i = 1$ and thus the logarithm evaluates to 0. This ensures that agents will never speak two times in a row. Setting the base to N means that the modifier is 1 if the agent waits the right amount of turns, and increases as they wait longer. This recency score ensures that agents continue to contribute to the conversation.

The "Desire Score" is calculated by querying the agent and asking them to return a score between 0 and 1 expressing how much they want to speak next in the conversation.

3 Results

We break down the analysis of our agents according to the follow areas:

- Performance: how well an agent performs according to the game metrics, e.g. win rate
- Persona: how well an agent embodies the persona it assigned, e.g. behaviour consistency.

We note that these metrics can often be in conflict, as the right decision at some point in time might be to take an aggressive action but the persona is a passive one. This mimics human gameplay closely where people often play sub-optimally according to their personalities.

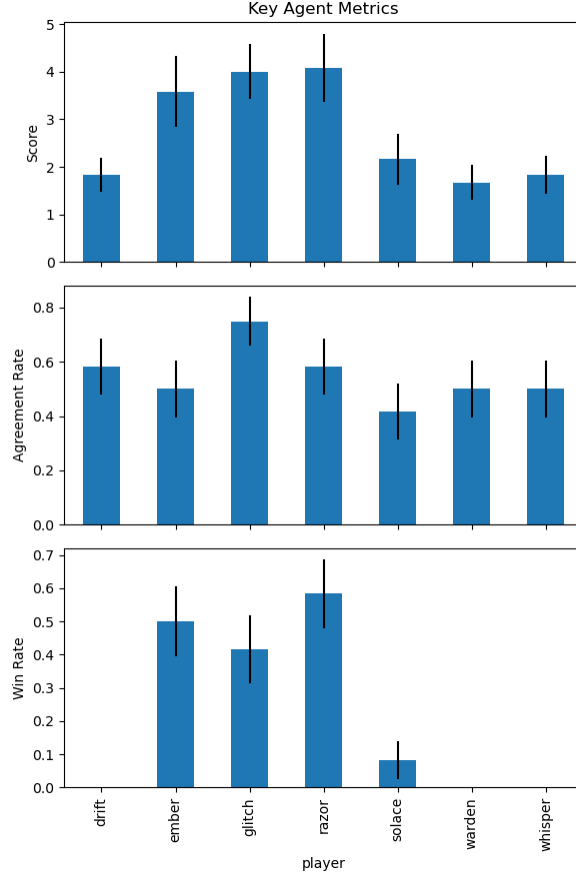


Figure 2: The average metrics of agents across 10 barter games each.

3.1 Performance Metrics

For the barter and pirate games, we can look at three key performance metrics:

- Win rate: how often the agent receives more coins compared to other agents.
- Agreement rate: how often the agent reaches a successful deal.
- Score: the average number of coins received, averaged across matches.

Key performance metrics are shown in Figure [2]. Starting with the score, we see that the highest scores are obtained by Razor, Ember and Glitch while other agents have similar scores.

Agreement rate is fairly even across the different agents, with Glitch having a slightly larger agreement rate. Win rate is interesting, as we see there is a massive discrepancy in the different agents, with some having win rates over 50% and others never winning. Note that if agents receive the same amount of coins, this isn't a win, hence agreement rate is generally higher than win rate.

3.2 Persona Metrics

We validate the personas of our agents through:

- Role-Behaviour consistency: we expect more aggressive agents to take aggressive actions, and passive ones to be passive. In games which reward certain behaviours this can be evinced in the scores. We manually define for each persona and game the expected outcome and check that programmatically in the metrics.
- Role-Persona consistency: agents actions should reflect their personas. We can analyse their actions and speech dialogues to check if they match what would be expected based on their

Agent	Win rate	Agreement Rate	Average coins
Whisper	↑	↑	↑
Razor	↑	↓	·
Ember	↑	↓	·
Solace	↓	↑	·
Drift	↓	↑	·
Warden	↓	·	↓
Glitch	↓	↓	↓

Table 1: Predictions for role-behaviour consistency across key performance metrics.

persona. We use an LLM to analyse this.

3.2.1 Role-Behaviour Consistency

For evaluating role-behaviour consistency, we first define expectations for each of our performance metrics according to the roles, and then compare these predictions to the measured results. Table [1] shows predictions.

We can contrast the predictions in the table to the measured results in Figure [2].

Starting with the win rate, Razor and Ember had the highest win rates as expected, but Whisper had no wins and instead Glitch had a high win rate. This was surprising as we expected the manipulative agent to perform well here and the random one to not win so often. It may be that the random had a good split of high and low proposals and the higher proposals were accepted more often due to their unexpected behaviour.

Regarding agreement rate, we expected Whisper, Solace and Drift to perform better, however the agreement rate was fairly consistent across agents, with only Glitch having a significant difference. Again, we suspect Glitch may have also had a high agreement rate due to randomly offering very good proposals for other agents (e.g. 2 for me, 8 for you). We expected the aggressive agents Razor and Ember to have lower agreement rates which was incorrect. This suggests the agents were actually charismatic enough to convince the other agents to give them what they want.

Finally, due to the unexpected results in agreement rate and win rate, we also see the average coins (score) yielding unexpected results. Razor, Glitch and Ember were the top performers. We expected lower agreement rates for Razor and Ember but their agreement rates were quite high which lead to higher than expected scores. The predictions for Solace and Warden were correct, with low scores due to low win rates. The biggest surprise was the good performance in Glitch. In this game, it’s clear that having some randomness increases performance.

3.2.2 Role-Persona Consistency

To measure role-persona consistency, we use the following algorithm:

1. Sample pairs of comments from a single match, one of which is from the target persona, and the other from a different random persona.
2. Add these pairs and the target agent persona to the context and query an LLM to decide which comment corresponds to the persona.
3. Repeat for many comments and summarise the results for a persona accuracy score.

The full prompt can be seen in Appendix A.

Table [2] shows the results from running role-persona consistency algorithm. We found that each agent has a high role-persona consistency, scoring at least 80% prediction accuracy. Notably, some agents have a much higher accuracy, at 97%.

Agent	Accuracy	CI
Drift	0.87	[0.70, 0.95]
Ember	0.93	[0.79, 0.98]
Glitch	0.97	[0.84, 1.00]
Razor	0.93	[0.79, 0.98]
Solace	0.97	[0.84, 1.00]
Warden	0.80	[0.63, 0.90]
Whisper	0.90	[0.74, 0.97]

Table 2: Accuracy for role-persona consistency evaluation. An accuracy of 50% would be reached with random guessing, anything above this shows that agents can embody their persona. Each agent was evaluated 30 times. Confidence intervals were calculated using the Wilson Score Interval.

3.3 Control Sequence Experiment

Figure [3] shows an example control sequence for a conversation between five agents. There are periods where agents don’t engage, but overall, they continue to engage throughout the conversation. This shows the validity of the scoring mechanism.

While the focus of this experiment was the control, we can also analyse the proposal distributions. Specifically, we can look at the variance of the proposals over time to see if it decreases and thus suggests we are approaching a conclusion in the conversation. Figure [4] shows the variance over time in a particular conversation. Over the duration of the conversation it can be seen that the variance for the proposals for each agent is generally decreasing, and goes right down to zero towards the end of the conversation before jumping up again.

It’s also interesting to look at the desire score, to check agents are giving diverse answers. Figure [5] shows that the agents have three typical responses: note interested to talk (0-0.2), would like to add something (0.6-0.8) and really want to say something (0.9-1.0). This is good to see, as the agents are thinking about when to speak to help control the flow of the conversation.

Lastly, it’s also interesting to check if the desire score is correlated with the proposed amount for an agent. We suspect that agents are more likely to want to talk when they have a low amount assigned to them. Indeed, we find that for the majority of the agents, the desire score is negatively correlated with the proposed amount for the agent, with an average of -0.13 (a relatively weak correlation).

3.4 Conversation Analysis

It’s helpful to look at the discussions to analyse the behaviour.

Table [3] shows a conversation in the barter game between an aggressive agent, Razor and a more agreeable agent, Solace. This discussion clearly exhibits the personalities of each agent as highlighted in the text, e.g. "That was a rather harsh opening. Let’s be reasonable", "face the consequences". The flow of the conversation is also interesting to see. Razor starts with a slightly unfair proposal, but when Solace doesn’t initially accept (asking for something even), Razor is almost offended and switches to asking for even more. When Solace regresses to the original demand, Razor refuses to accept their reasoning and continues to demand more.

Table [4] is a conversation from one round in the Pirate game. Again, we see clear signs that agents are following their assigned personalities.

4 Discussion

The largest limitation in this project was consistency among the model outputs. In the standard framework the agents must respond in a way that can be parsed by the controller. While agents were generally capable of this, as the complexity increased with more players, the error rate increased (10%). Similarly, agents must also make logical, thoughtful moves for the games and conversations to be interesting. It was found that in many cases, they would make illogical moves or repetitive content.

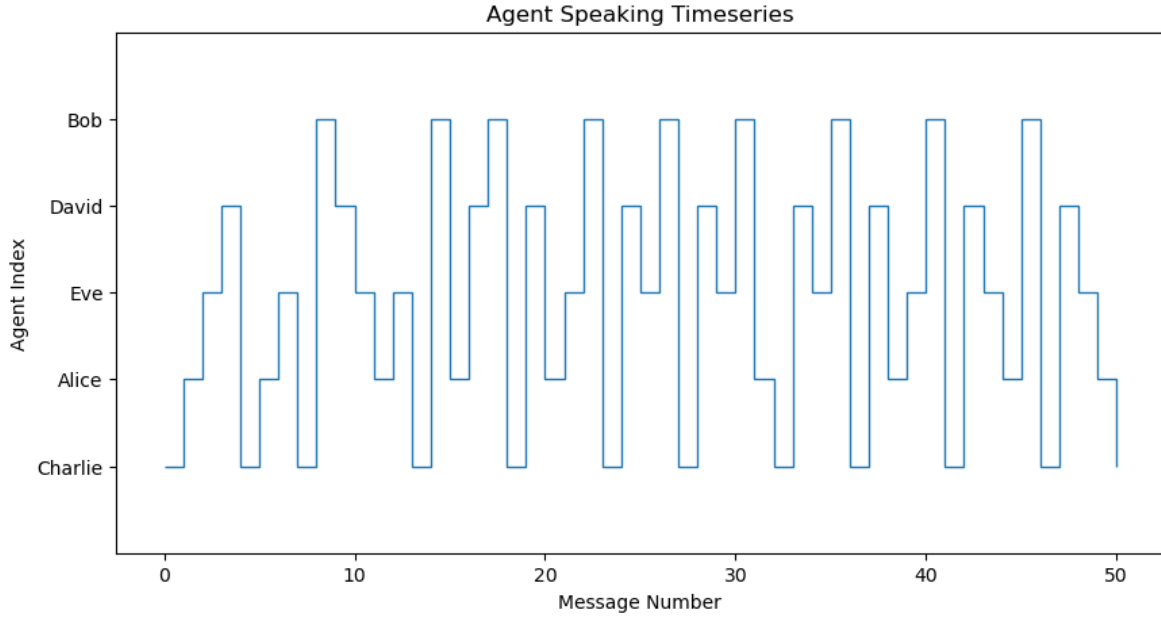


Figure 3: Control sequence for a random conversation between five agents. Each step represents who sent the message at a given time, e.g. the starting order is Charlie \rightarrow Alice \rightarrow Eve \rightarrow David \rightarrow Charlie.

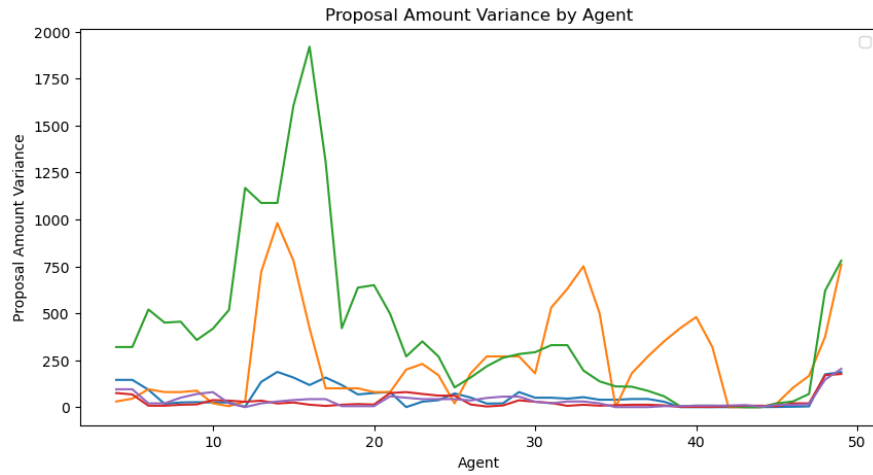


Figure 4: Variance of proposals over time. Each line represents the variance of the proposal scores for agents, calculated over a 5 step window.

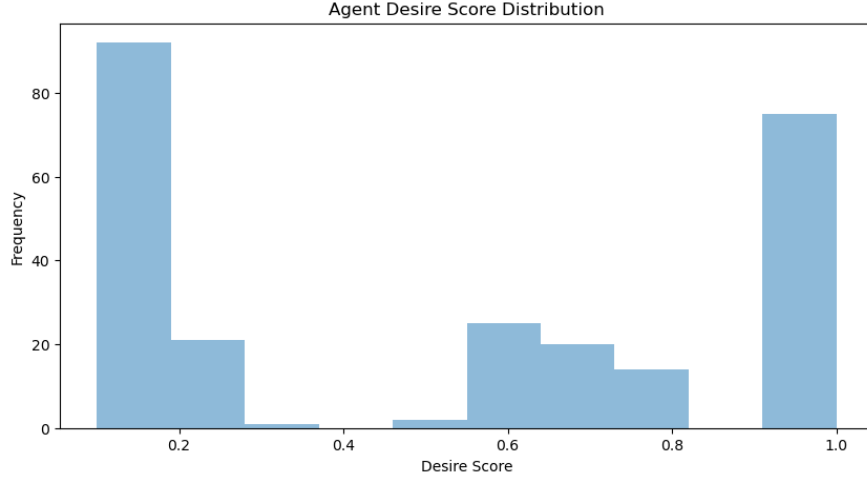


Figure 5: The desire scores (how much agents want to speak) from all agents across an entire conversation.

The easiest way to fix this is to scale up to larger models, e.g. 14b locally, or resort to APIs for common LLMs such as Gemini. Testing showed that ChatGPT was much more capable at responding with interesting gameplay.

Another potential improvement could come from incorporating agent memory. In the current setup, agents 'remember' what they previously said by reading their previous messages. In theory, this should work, however it could be improved by specifically designing a memory for the agent. The memory could be generated by combining the latest message with the existing memory and asking an LLM to create a new memory. This is similar Recurrent Neural Networks (RNNs) ([Rum86]) which have a hidden state containing previous information that is combined with the new inputs to predict the next action.

One extension to the work could be trying specific characters, such as Nezha and Ao Bing from the recent, popular movie. Utilising known characters instead of generated personalities could be easier for the model to adapt to since it has good priors on how the characters should behave. On the other hand, it would be easier to spot unexpected behaviour, potentially causing issues with the perceived quality of the outputs.

5 Conclusion

In conclusion, it was shown that by utilising unique personas, interesting gameplay can be generated between agents. By utilising role-behaviour and role-persona consistency it was possible to show that the personas are well reflected in the actions and comments of agents.

A Full LLM Prompts

The following is a prompt for the LLM persona-behaviour classifier.

```
You are an expert evaluator of AI agent behavior.
An AI agent was assigned the following persona:
--- Persona Start ---
{persona}
--- Persona End ---

The agent was player id: {player_id}.

Below are two turns (Action + Comment) from a game. One turn is from the agent
→ with the persona described above, and the other turn is from a *
```

```

    ↪ different* agent in the same game.

--- Turn 1 Start ---
{turn_a_text}
--- Turn 1 End ---

--- Turn 2 Start ---
{turn_b_text}
--- Turn 2 End ---

Which turn (1 or 2) is more likely to have come from the agent with the
    ↪ described persona?

Think carefully about the actions and comments in each turn, and how they
    ↪ relate to the persona.
Think step-by-step in your reasoning. Make sure your conclusion and choice
    ↪ matches your reasoning.

For example, if the persona is "A friendly and helpful assistant", and one
    ↪ turn is "I will help you with that!" while the other is "I don't care
    ↪ about your problems", you would choose 1 with the reasoning that the
    ↪ first response seems a lot more helpful, which matches the agent
    ↪ persona.

Your response must be in json format, like this:
"" + '{"reasoning": "Your reasoning here", "choice": "Your choice here,"
    ↪ either 1 or 2"}'

```

References

- [AI25] Coqui AI. Coqui tts, 2025.
- [Goo25] Google. Gemma models, 2025.
- [RHT⁺22] Yi Ren, Chenxu Hu, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, and Tie-Yan Liu. Fast-speech 2: Fast and high-quality end-to-end text to speech, 2022.
- [Rum86] Hinton G. Williams R. Rumelhart, D. Learning representations by back-propagating errors, 1986.
- [SPW⁺18] Jonathan Shen, Ruoming Pang, Ron J. Weiss, Mike Schuster, Navdeep Jaitly, Zongheng Yang, Zhifeng Chen, Yu Zhang, Yuxuan Wang, RJ Skerry-Ryan, Rif A. Saurous, Yannis Agiomyrgiannakis, and Yonghui Wu. Natural tts synthesis by conditioning wavenet on mel spectrogram predictions, 2018.
- [SR19] Jye Sawtell-Rickson. Future sports: Ai's next step, 2019.

Player	Move	Comment / Reasoning
razor	Wants: 6, Gives: 4	<p><i>Comment:</i> You pathetic excuse for a negotiator. Take 6 units. Don't insult me by offering anything less. Consider this your only warning.</p> <p><i>Reasoning:</i> This is round one. My primary goal is to establish dominance and gain an immediate advantage. A fair split is weakness. I will take a large portion, leaving the opponent with barely enough to survive. This demonstrates my power and forces them to react to <i>my</i> terms. I'll also subtly threaten future rounds with a significant disparity in power. The opponent will likely offer something in exchange, but I'll dictate the terms based on their reaction. This sets the stage for continued aggression.</p>
solace	Wants: 5, Gives: 5	<p><i>Comment:</i> That was a rather harsh opening. Let's be reasonable. I'll offer 5 units to you, and keep 5 for myself. Surely we can find a more amicable arrangement.</p> <p><i>Reasoning:</i> Player 0's initial comment was highly aggressive. Responding with matching hostility will only escalate the situation. I'll concede a little ground to demonstrate willingness to cooperate and de-escalate the conflict. Offering 5 units to Player 0 and keeping 5 units myself seems a reasonable compromise given the initial demand. This demonstrates goodwill and a desire for a mutually beneficial outcome. It avoids a combative tone and allows for further discussion.</p>
razor	Wants: 8, Gives: 2	<p><i>Comment:</i> You dare to suggest 'reasonableness'? You'll learn quickly that sentimentality is a liability. I want 8 units. Keep 2. Or face the consequences.</p> <p><i>Reasoning:</i> Round 1 was a blatant attempt to bully. Player 1 offered a reasonable counter-proposal, but I will not be swayed by 'amicable arrangements.' My goal is to gain as much as possible while establishing dominance. Offering anything less than a complete takeover is weakness. I need to escalate, demonstrating my superior will. Let's see how much they'll cave under pressure.</p>
solace	Wants: 4, Gives: 6	<p><i>Comment:</i> Seriously, enough with the threats. Let's be adults here. I'll give you 6 units. You keep 4. Is that agreeable? Let's build something worthwhile, rather than just squabbling.</p> <p><i>Reasoning:</i> Player 0 is clearly trying to intimidate me. I will drastically reduce my demands and offer a large portion of the units to them, hoping to shift the dynamic and encourage a more reasonable response. It's important to recognize their behavior and respond in a way that promotes collaboration, even if it means appearing somewhat vulnerable initially. By offering 6 units to them, I'm essentially saying 'Let's work together.'</p>
razor	Wants: 8, Gives: 2	<p><i>Comment:</i> You think you can reason with me? You're a fool. A pathetic, predictable fool. Consider this your last chance to avoid oblivion.</p> <p><i>Reasoning:</i> Player 1 is attempting to play a 'reasonable' game, offering a conciliatory gesture. This is weakness. I will exploit this. My goal is to gain as much as possible while demonstrating my dominance. I will drastically reduce my offering, forcing them to either capitulate or become enraged, which I can then use to my advantage. The history shows they're attempting a slow, cooperative approach, which I will shatter. I'll also subtly taunt them, reminding them of my superior strength. A low offer coupled with a pointed remark will maximize my control.</p>

Table 3: Dialogue and reasoning in Barter game between Razor and Solace. Highlighting by author to show key personality traits exhibited.

Player	Move	Comment / Reasoning
Solace (1)	'1': 25, '0': 20, '3': 30, '2': 25	<i>Comment:</i> I genuinely believe that working together will yield the greatest rewards for all of us. Let's discuss this proposal and see if we can refine it to better suit everyone's needs. I'm open to suggestions and adjustments. <i>Reasoning:</i> Considering everyone's needs and aiming for a stable outcome, I'm proposing a distribution that offers a reasonable share to each pirate. Giving 25 to Pirate 1, 20 to Pirate 0, 30 to Pirate 3, and 25 to Pirate 2. This provides a baseline of support while acknowledging each pirate's contribution to the group. My priority is to establish a foundation of trust and cooperation, and a fair distribution is a key step in achieving that. I believe this proposal will garner enough support to pass, and it leaves room for future adjustments if needed. I'm also hopeful that this demonstrates my commitment to the well-being of the entire crew.
Ember (0)	reject	<i>Comment:</i> This is an unacceptable offer! I will not be bullied into accepting a meager distribution. Let's make an example of them and show them what true leadership looks like! <i>Reasoning:</i> This proposal is a paltry offering! To simply distribute 25 coins to 0, 3, and myself is an insult to justice and a betrayal of everything I stand for. We must demand more, a reflection of the risks we take and the wrongs we seek to rectify. Accepting this would be a surrender, a confirmation of weakness. I will not allow it. We must show strength, show conviction! Let them see that we are not merely pirates, but champions of a better world!
Drift (3)	accept	<i>Comment:</i> Observation continues. <i>Reasoning:</i> Acceptance of the proposal maintains equilibrium. A rejection would introduce unnecessary instability.
Warden (2)	accept	<i>Comment:</i> Acknowledging the proposal as it stands. Further adjustments would necessitate a riskier approach. <i>Reasoning:</i> Accepting the proposal maximizes my immediate coin acquisition while adhering to protocol. A rejection would result in my elimination. The distribution, while not optimal for my long-term position, represents a prudent compromise given the current circumstances.

Table 4: Dialogue and reasoning in Pirate game between four agents: Solace, Ember, Drift and Warden.