Audio

Policy Model

LALM trainable

Text prompt

with 3 types of Reasoning Format

| IR | EUR | ESR |

Sampling Action Groups

```
<think>...</think><answer>...</answer>
<think>...</think><answer>...</answer>
        ...                ...
<think>...</think><answer>...</answer>
```

Updating Policy Model

KL Divergence

Reference Model **LALM** pretrained

Reward Evaluation

Emotion Similarity-Weighted Reward

Explicit Structured Reasoning Reward

Plutchik's wheel of emotions

Joy Trust Fear Surprise Neutral

Sadness Disgust Anger Interest

Emotion-State-Transition Matrix

calculate

Predicted Emotion

Groundtruth

Emotion Similarity-Weighted Reward

$$\mathcal{R}_{\text{ESWR}} = \begin{cases} 1 & S(\hat{y}, y) = 1 \\ \alpha \cdot S(y_{\text{pred}}, y_{\text{true}}) & S(\hat{y}, y) > \gamma \\ 0 & S(\hat{y}, y) \leq \gamma \end{cases}$$

Implicit Reasoning

Please think step by step. Output the final answer in **<answer> </answer>**.

(IR)

Explicit Unstructured Reasoning

Please think step by step. Output the thinking process in **<think></think>** and final answer in **<answer> </answer>**.

(EUR)

Explicit Structured Reasoning

Please think step by step. Your reply should strictly follow the following format: **<think>** **<content>**recognize and analyze the speaker's speech content) **</content>** **<acoustics>** (Secondly analyze the speaker's acoustic features) **</acoustics>** **<summary>** (Next summary the all views) **</summary>** **</think><answer>** (Finally give the selected emotion answer) **</answer>** **(ESR)**

Explicit Structured Reasoning Reward

$$\mathcal{R}_{\text{ESR}} = \begin{cases} 1 & \text{if the format matches } Q_{\text{ESR}} \\ 0 & \text{otherwise.} \end{cases}$$