



EmoDynamix: Emotional Support Dialogue Strategy Prediction by Modelling Mixed Emotions and Discourse Dynamics

Chenwei Wan, Matthieu Labeau, Chloé Clavel

Motivation

- **End-to-end chatbots** based on LLMs tend to overly rely on certain dialog strategies, and lack social cognitive ability, transparency and controllability.
- We decouple strategy prediction from generation, and present a **dialog strategy prediction framework** using explicit cognitive modeling.

Task

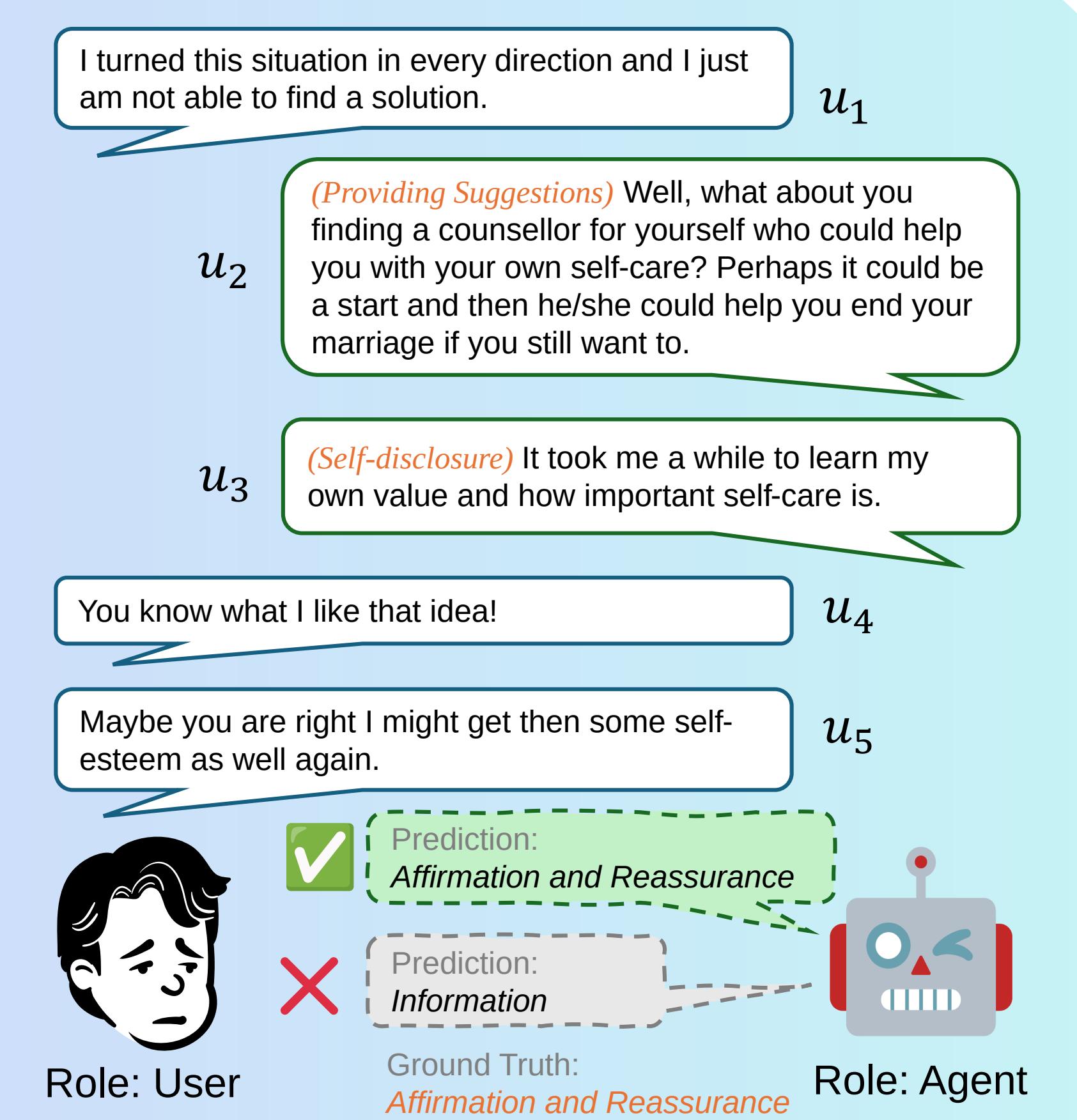
Input: Dialog History

- Dialogue context
- The past strategies applied by the agent

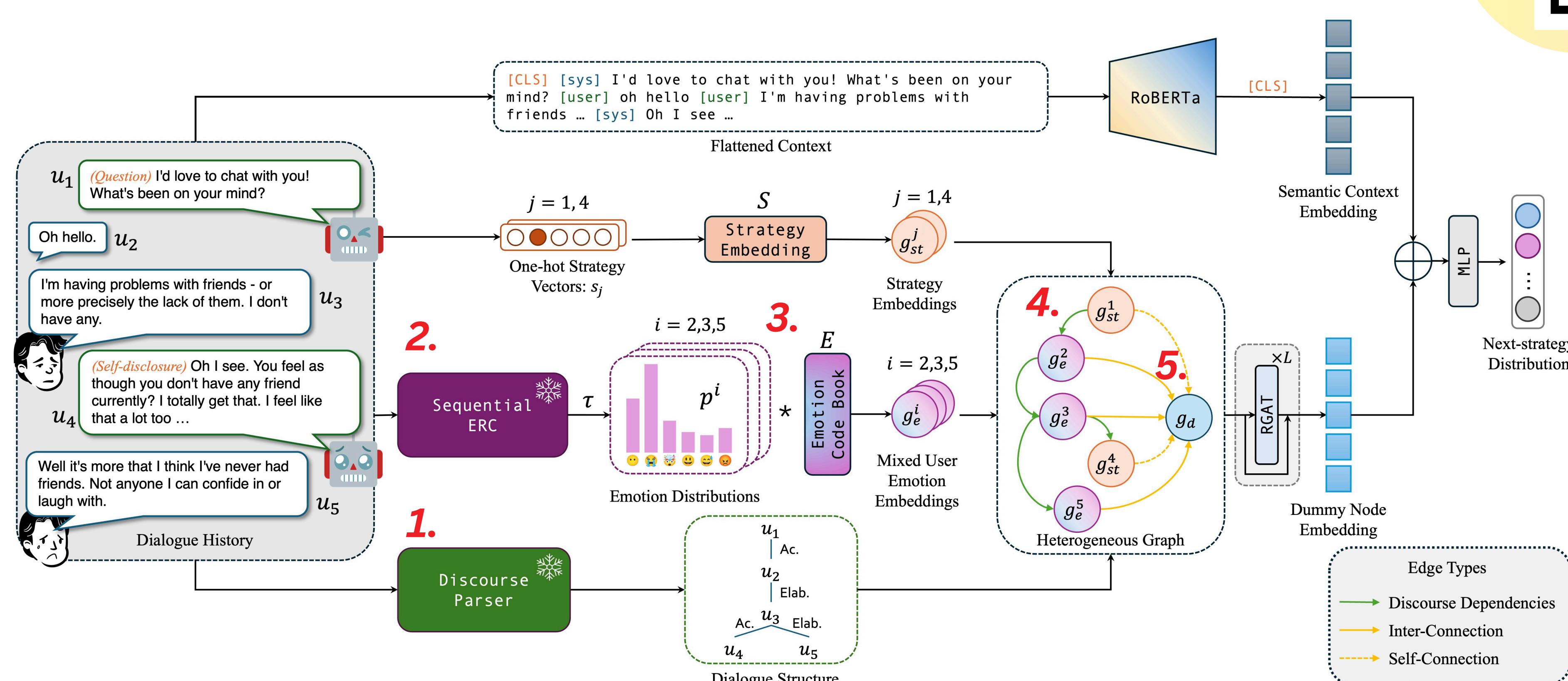
Output: Next dialog strategy for controlling the generation of the agent's next response
Label: one ground-truth strategy per dialog snippet

Metrics:

- Proficiency: Weighted & Macro F1
- Preference Bias score (Kang et al., 2024)



Methodology



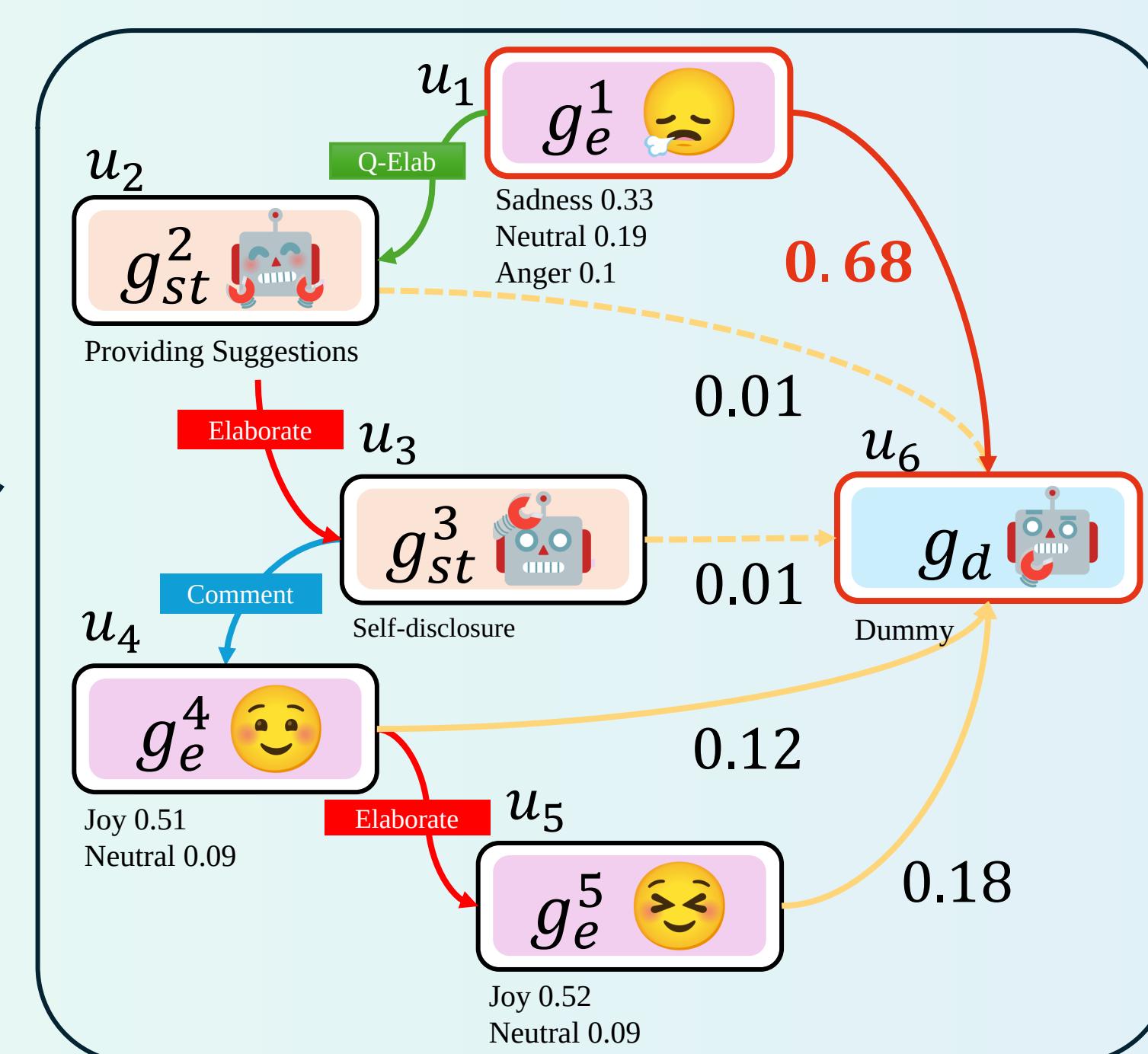
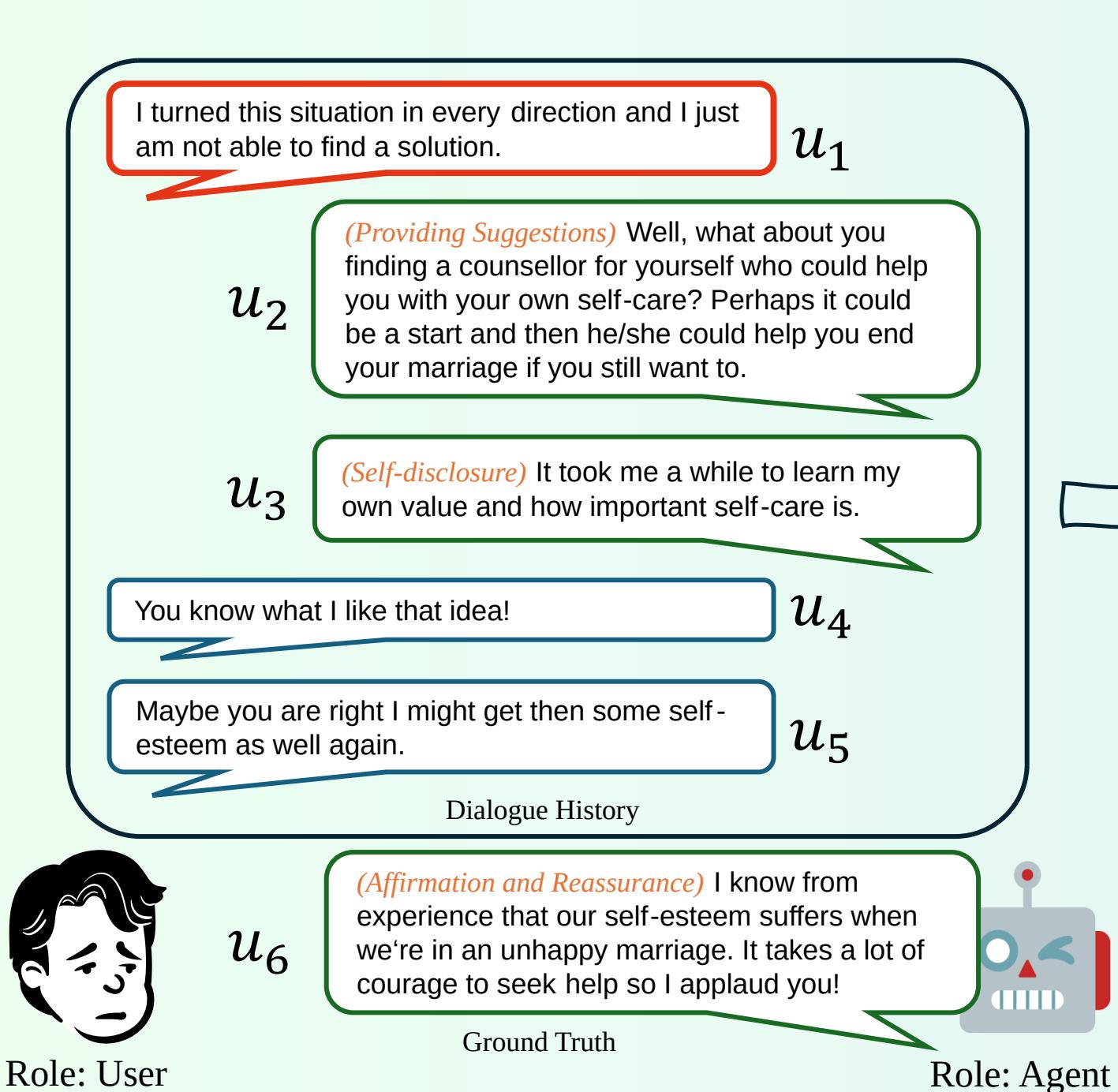
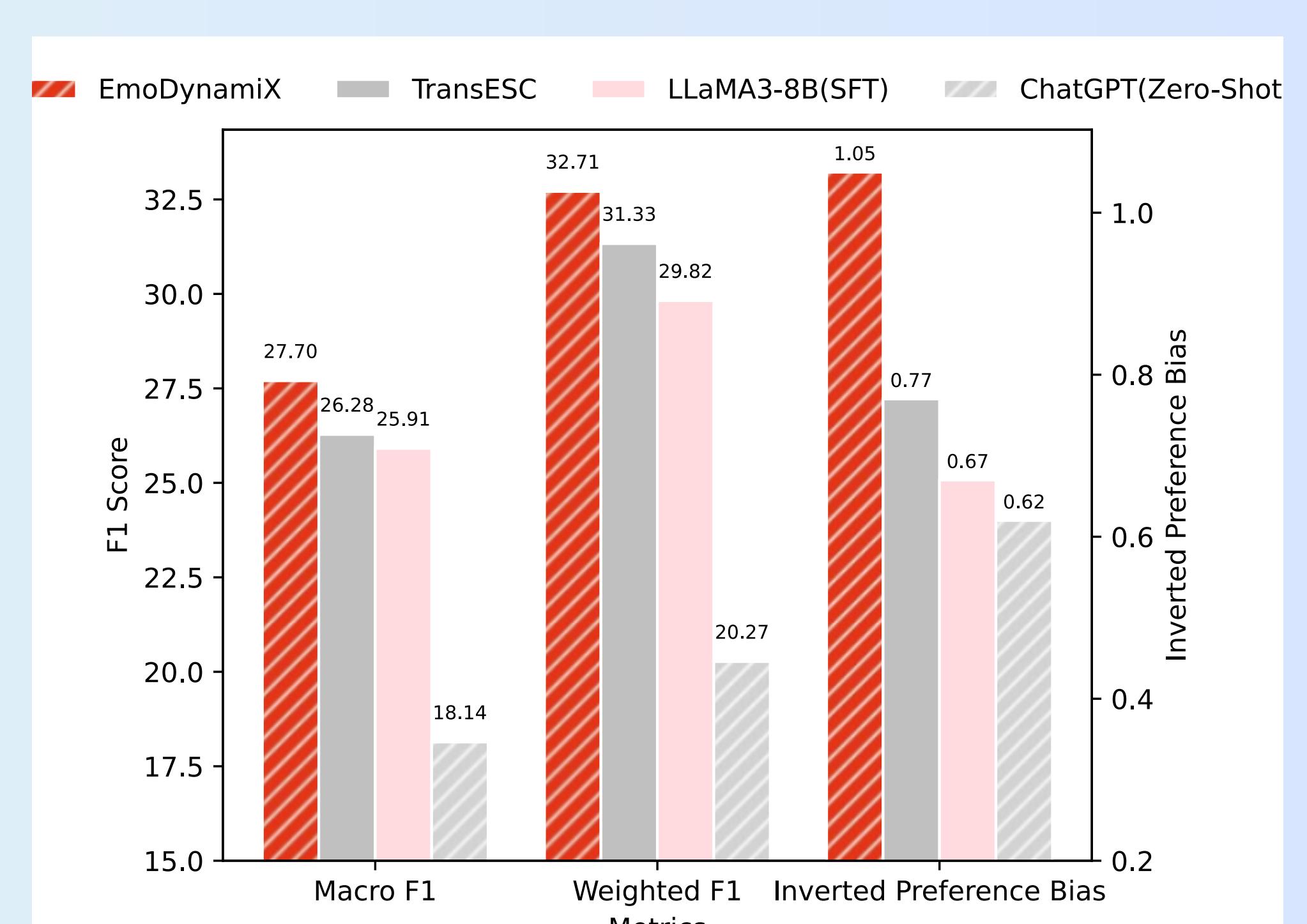
1. **Pre-trained discourse parser** for mining the dependencies between speaker turns.
2. **Pre-trained emotion recognition model** for mental state inference.
3. **Fine-grained mental states** are represented with mixed emotions: **temperature-controlled emotion distributions**
4. **Graph neural network** for representing the speaker-aware conversational dynamics.
5. **Dummy node** for information aggregation.

Experiments & Case Study

EmoDynamix outperforms:

- **SOTA ESC frameworks**: TransESC (Zhao et al., 2023), etc.
- **Finetuning SOTA LLMs**: LLaMA3-8B (Meta, 2024), etc.
- **Prompting SOTA LLMs**: ChatGPT (OpenAI, 2023), etc.

Across all metrics on **ESConv** (Liu et al., 2021) and **AnnoMI** (Wu et al., 2022).



- Extra transparency are provided with back-tracing attention weights.
- Mixed-emotion module can model a large set of subtle emotional expressions by combining primary emotions.

A diagram shows a user utterance "I turned this situation in every direction and I just am not able to find a solution." being processed by the system to infer an emotion. The inferred emotion is g_e^1 (Sadness 0.33, Neutral 0.19, Anger 0.1). A note says: "Emotion here can be inferred as **Frustration**, which does not belong to common emotion categories in ERC datasets." Another note says: "Primary emotion categories from DailyDialog (Li et al., 2023)".