Infusing Emotions into Task-oriented Dialogue Systems: Understanding, Management, and Generation

Shutong Feng*, **Hsien-chin Lin***, **Christian Geishauser***, **Nurul Lubis**, **Carel van Niekerk**, **Michael Heck**, **Benjamin Ruppik**, **Renato Vukovic**, **Milica Gašić**
Heinrich Heine University Düsseldorf, Germany
{fengs,linh,geishaus,lubis,niekerk,heckmi,ruppik,revuk100,gasic}@hhu.de

Abstract

Emotions are indispensable in human communication, but are often overlooked in task-oriented dialogue (ToD) modelling, where the task success is the primary focus. While existing works have explored user emotions or similar concepts in some ToD tasks, none has so far included emotion modelling into a fully-fledged ToD system nor conducted interaction with human or simulated users. In this work, we incorporate emotion into the complete ToD processing loop, involving understanding, management, and generation. To this end, we extend the EmoWOZ dataset (Feng et al., 2022) with system affective behaviour labels. Through interactive experimentation involving both simulated and human users, we demonstrate that our proposed framework significantly enhances the user's emotional experience as well as the task success.

1 Introduction

In recent years, conversational artificial intelligence (AI) has become increasingly prevalent in various domains, providing users with interactive and personalised experiences. Emotions play a crucial role in human communication and can influence the way individuals perceive, process, and react to information (Ekman, 1992). Consequently, incorporating emotions into conversational AI has emerged as a promising avenue for improving user experience and creating more human-like interactions (Picard, 2000).

Task-oriented dialogue (ToD) systems, an important genre of conversational AI, are designed to assist users in fulfilling specific tasks or queries. In contrast to chit-chat or open-domain dialogue systems, which focus on creating engaging and entertaining conversations, ToD systems interact with users in a more structured way with a clear objective under specific domains (Jurafsky and Martin, 2009). While significant advancements have been made in natural language processing and ToD systems, there remains a critical challenge in creating systems that can understand and respond to not only the informational needs of users but also their emotional states.

In ToD, emotion is centred around the user goal, making it more contextual and subtle (Feng et al., 2022). A recent study has shown that the valence of user emotion in ToD positively correlates with dialogue success (Lin et al., 2023). This observation aligns with a number of emotional theories. For example, the appraisal theory of emotion argues that emotion is the result of our evaluation of a situation (Arnold, 1960; Lazarus, 1966). In relation to a ToD user goal, it is straightforward to see how task fulfilment would lead to positive emotions and failures to negative ones. Similarly, the Ortony-Clore-Collins (OCC) model of emotion states that emotion is the result of elicitation by events, agents, and objects (Ortony et al., 1988). Feng et al. (2022) have drawn the connection between the OCC model and user emotions in ToD. Therefore, besides inferring emotional states from dialogue utterances, an agent also needs to reason about emotion-generating situations and to utilise this information to achieve dialogue success.

The integration of emotion into the full ToD pipeline has been a long-standing interest (Bui et al., 2010; Ren et al., 2015). Yet, early works explored analytical solutions in constrained set-ups, which hindered their applications in more complicated scenarios. Recently, a number of resources emerged for studying user affect in ToDs, e.g. emotion, sentiment, or satisfaction (Mendonca et al., 2023; Feng et al., 2022). This has motivated efforts to model user emotion via data-driven approaches, such as emotional user simulation (Lin et al., 2023) and user emotion recognition (Feng et al., 2023a; Stricker and Paroubek, 2024). However, to the best of our knowledge, no work so far has com-

*These authors contributed equally to this work.

Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue, pages 699–717
September 18–20, 2024.

bined these emotional aspects into a fully-fledged dialogue system and an interactive pipeline where emotions play a role in understanding, generation, as well as management of the conversation.

To achieve this, we need to endow the dialogue system with the ability to respond with an affective behaviour, closing the emotional loop between the user and the system in ToDs. Towards this goal, we make the following contributions:

- We extend EmoWOZ, a large-scale ToD dataset for user emotions (Feng et al., 2022), with annotations for *affective conduct* in 71k system utterances. To the best of our knowledge, this is the first large-scale and open-source corpus dedicated to the system's affective behaviour in ToDs.
- We incorporate emotion in the complete ToD interaction loop for understanding, management, and generation by building a modular system around an emotion-*aware* and emotion-*expressive* policy. We also build an emotional LLM-based end-to-end ToD system that involves emotion in understanding and generation.
- For our modular system, we train our dialogue policy via reinforcement learning (RL) on the natural language level, leveraging emotions and task success as reward signals. We train the end-to-end system on our newly collected dataset via supervised learning (SL). For both systems, we show through interactive evaluation that emotion in the ToD loop can enhance user's emotional experience as well as the task success. This highlights the importance of modelling emotions in ToDs.

2 Related Work

In this section, we discuss related works on incorporating emotion in each stage of ToD pipeline: understanding, management, and generation. These stages are modelled explicitly with multiple models in modular systems and implicitly with a unified model in end-to-end systems (Hosseini-Asl et al., 2020; Stricker and Paroubek, 2024).

2.1 Understanding User Emotion

Modular ToD systems rely on natural language understanding (NLU) and dialogue state tracking (DST) modules to translate and accumulate semantic concepts related to user goals. Typically, these semantic concepts are strictly limited to those defined in the ontology, i.e. domains, slots, and values the system can talk about.

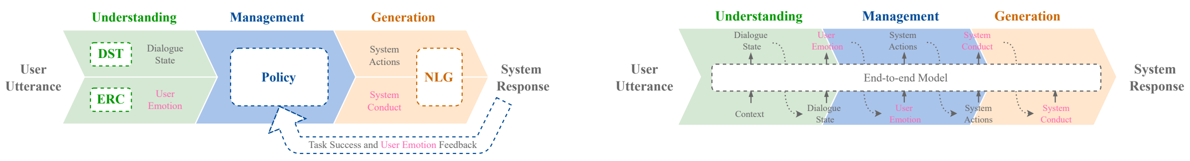
Given its potential as an important piece of information for the system's subsequent decision-making, emotion can be considered as part of the dialogue state. Feng et al. (2022) showed that multi-task training a DST model for emotion recognition simultaneously improves its joint goal accuracy, suggesting the complementarity between DST and emotion recognition in conversation (ERC). Recently, Stricker and Paroubek (2024) modelled user emotion as an intermediate task in end-to-end ToD systems and improved overall system performance. Standalone ERC models dedicated to ToDs (Li et al., 2023; Feng et al., 2023b) can be used in modular systems in parallel with any DST to extend the dialogue state with user emotions.

2.2 Dialogue Management with Emotion Feedback

In ToD, one way to train the dialogue policy is via RL to maximise task success, indicated at the end of the dialogue based on user goal fulfilment (Levin and Pieraccini, 1997; Kwan et al., 2023). Since user emotion is highly associated with task success (Lin et al., 2023), it is intuitive to leverage user emotion during the dialogue for providing more dense and diverse reward signals. Bui et al. (2010) incorporate user emotion into the policy state by modelling affective dialogue management through a factored partially-observable Markov decision process (POMDP) and analytically find an optimal policy. This is however neither feasible for larger problems, nor has this been integrated in interactive set-ups. Zhang et al. (2021) addressed the delayed reward problem in dialogue policy learning with a predefined emotion-based turn-level reward. Zhu et al. (2024) consider the difference between the user's positive emotion intensity and the next turn's emotion utility value for top-k action selection. We take a step further by incorporating emotion in policy state *and* reward function. We then leverage emotion in RL to find optimal semantic actions *and* affective expression of the system, which has not been explored before.

2.3 Generating Affective Response

The natural language generation (NLG) module in ToD systems realises semantic actions from the policy into natural language. Traditionally, ToD NLG focuses on translating task-related semantic



(a) Modular Systems (b) End-to-end Systems

Figure 1: Infusing emotions into modular and end-to-end ToD systems.

actions and overlooks other aspects of system responses such as emotion. There have been efforts to create datasets which help enrich ToD system responses with chit-chat (Sun et al., 2021; Chen et al., 2022; Young et al., 2021). Stricker and Paroubek (2024) attempted to refine end-to-end system output with a large language model (LLM) under a chain-of-thought framework to enhance simulated system empathy. Different from previous works, we aim to enrich system response with the subtle affective conduct jointly with dialogue actions in a fully controllable approach.

2.4 Simulating User Emotional Behaviour

User simulators (USs) simulate user behaviour in ToDs. Although they are not a part of the system, they play essential roles in training dialogue policy via RL and serving as an efficient evaluation platform for dialogue policy (Eckert et al., 1997). Most existing USs focus on modelling user's behaviour in terms of semantic actions and natural language by taking system semantic actions (Kreyssig et al., 2018; Lin et al., 2021, 2022). Zhang et al. (2021) built a US that additionally incorporated handcrafted emotion transitions in different situations. Kim and Lipani (2022) used a data-driven approach and simulated satisfaction levels along with the intent and the utterance. Lin et al. (2023) further proposed data-driven EmoUS to model more nuanced user emotions with enhanced controllability via user persona settings. This motivates us to move one step further to capture more fine-grained affective expressions of the system from natural language response directly.

3 EmoWOZ 2.0: A Fully Emotion-annotated ToD Dataset

To study emotion in real-world interactions between users and human operators in the ToD setting, we extend EmoWOZ (Feng et al., 2022) by further annotating the *affective behaviour* of the system, which is acted by human-beings. We call this dataset with extended labels *EmoWOZ 2.0.* [*]

In ToDs, the user and the system play different roles. Users may express a wide range of emotions during interactions based on their goals and experiences with the operator. The system is responsible for managing and facilitating the conversation and is supposed to behave professionally and politely to achieve the goal. Therefore, it is necessary to consider different sets of affective behaviours in the user and the system respectively. We refer to the concept of the operator's affective behaviour as **affective conduct**, or **conduct** for short.

Annotation Scheme According to studies on customer satisfaction in business (Grandey et al., 2011), competent operators in ToD try to guide user emotion towards positive valence by making use of subtle emotion in their response while providing correct information. By considering the set of user emotions in EmoWOZ and the OCC emotion model (detailed justification in Appendix A.1), we arrive at five affective conduct classes:

- **Neutral:** the operator does not explicitly make use of any affective conduct.
- **Compassionate:** the operator is sympathetic about user's situation, usually in response to a fearful/disappointed user in an unpleasant situation.
- **Apologetic:** the operator apologises for their mistake, usually in response to a dissatisfied user.
- **Enthusiastic:** the operator is feeling happy for the user or showing extra eagerness to help. This conduct takes place usually in response to a neutral or excited user.
- **Appreciative:** the operator acknowledges the – at least partial – task success, usually signalled as user's satisfaction.

Annotation Set-up We annotated the conduct for all operator utterances in the MultiWOZ subset of

[*]EmoWOZ 2.0 is released under CC By 4.0 NC license, following the original EmoWOZ release. The dataset can be found at `https://gitlab.cs.uni-duesseldorf.de/general/dsml/emowoz-2.0-public/`

EmoWOZ. Machine-generated system responses in the DialMAGE subset came from a template NLG, which we considered to have neutral conduct because those templates aimed to express actions concisely rather than conveying emotions by design.

We followed the data collection and quality assurance set-up of EmoWOZ and conducted the annotation via the Amazon Mechanical Turk platform. Details and an illustration of the annotation interface can be found in Appendix A.2.

Annotation Quality Each utterance has been annotated by at least three annotators. The inter-annotator agreement as measured with Fleiss' Kappa is 0.647, suggesting substantial inter-annotator agreement. The annotator confusion matrix and label distribution can be found in Appendix A.3 and A.4, respectively.

4 Infusing Emotions into ToD Systems

We propose to incorporate emotion into the full interactive ToD pipeline, which is primarily comprised of three stages: understanding, management, and generation. We aim for understanding to accurately recognise the user's emotion in addition to the task-centred dialogue state. For dialogue management, we make use of emotion for optimal action selection. Lastly, we additionally condition the natural language generation on the system conduct to generate more diverse and emotion-aware responses. These are realised in each modular system component individually (Section 4.1 to 4.3) and as intermediate tasks in the unified model in end-to-end systems (Section 4.4). [*]

4.1 Expanding Dialogue State with Emotion

In our modular system, we use an ERC model in parallel with a DST model. This allows a flexible selection of DST and the associated ontology. The inferred user emotion is appended to the dialogue state.

For ERC, we use the ContextBERT-ERToD model (Feng et al., 2023a) as our user emotion recognition front-end because of its good ERC ability and fast inference. It is a BERT-based classification model (Devlin et al., 2019) that considers dialogue context and state in addition to the user utterance. It reports a weighted F1 score of 83.9% for emotions excluding neutral.

For DST, we use the SetSUMBT model (van Niekerk et al., 2021). This model, based on the RoBERTa language model (Liu et al., 2019) and a recurrent context tracker adopts a picklist approach to DST. Specifically, we employ the `Ensemble-Distribution-Distilled` variant of SetSUMBT, a refined version that distils knowledge from an ensemble of models. This version reports a joint goal accuracy of 51.22% on MultiWOZ. The architectural design of SetSUBMT also allows transferability to new domains, and such an ability has been exemplified with a transformer-based dialogue policy under a continual learning set-up (Geishauser et al., 2024).

4.2 Emotion-aware Dialogue Policy

For dialogue management in the modular system, we build a dialogue policy that considers the user emotion in the input and produces an emotion-augmented system output. We utilise the dynamic dialogue policy transformer (DDPT) architecture (Geishauser et al., 2022) since it was built for optimising dialogue policies that require extendable input and output, which facilitate the adaptation to new domains and ontologies. The dialogue policy leverages emotions in three ways: considering user emotion in the input, generating system affective conduct in the output, and considering user emotion in the reward for RL.

Emotion Input and Output The user emotion, as a part of the dialogue state, is incorporated into the dialogue state through embedding the perceived user emotion with RoBERTa. For semantic action selection, DDPT produces a sequence of domain-intent-slot triplets auto-regressively through its transformer decoder, e.g. `restaurant-inform-phone`, `restaurant-request-food`, until a stop token is generated. In order to predict *emotional* system conduct, after DDPT outputs the semantic actions, we decode the sequence for one more step to generate the system conduct action, considering the perceived user emotion from the dialogue state.

Emotion Augmented Reward We incorporate user emotion into the reward for RL by considering the associated sentiment. More specifically, we define $c(\text{satisfied}) = 1$, $c(\text{dissatisfied}) = c(\text{abusive}) = -1$, $c(\text{neutral}) = 0$. For the remaining user emotions that are not elicited by the

[*]The code of pipeline systems, end-to-end systems, and the user simulator can be found at https://gitlab.cs.uni-duesseldorf.de/general/dsml/emoloop-public/

system, we set $c(\text{emotion}) = 0$. For any emotion e, we multiply $c(e)$ by a hyperparameter β to weight the influence of emotion in the reward. Note that utilizing $\beta \cdot c(e)$ directly could encourage the dialogue policy to produce long dialogues with unnecessary turns as long as they produce positive user sentiment. In order to prevent this, we shift $\beta \cdot c(e)$ such that it is at most 0 by defining the emotion reward for an emotion e as $r_{\text{emo}}(e) = \beta \cdot c(e) - \beta$.

The emotional reward is combined with the standard reward r_{task} in dialogue policy learning that equals -1 in every non-terminating turn for encouraging efficiency and either $-T$ or $2T$ for dialogue failure or success, where T denotes the maximum permitted number of turns. The final reward is thus given by $r = r_{\text{task}} + r_{\text{emo}}$. We refer to this policy with expanded dialogue state input, expanded dialogue action output, and emotion reward as **EmoDDPT**.

4.3 Expressing Emotion in Response

Our modular system NLG was built based on the BART model (Lewis et al., 2020). We followed existing works to formulate the ToD NLG problem as a sequence-to-sequence task (Peng et al., 2020; Zhu et al., 2023) where the input is a sequence containing semantic concepts in textual form (e.g. tuples of [intent, domain, slot, value]), and the output is natural language conveying the semantic meaning. Our model input consists of the user utterance, system semantic actions, and the system conduct. We refer to our system NLG as **SEC-BART**: a both **s**emantically and **e**motionally **c**onditioned BART. In our ablation study, we used **SC-BART**, the version that is only conditioned on the semantic actions in the non-emotional ToD pipeline.

On MultiWOZ, SEC-BART achieves a BLEU score of 34.9 and a slot error rate of 3.6%, comparable to existing SOTAs (Peng et al., 2020; Zhu et al., 2023). Details of model training and performance can be found in Appendix C.

4.4 Emotional End-to-end System

We follow the work of Stricker and Paroubek (2024), where ERC is added as an intermediate task in the end-to-end ToD modelling, i.e. emotion is incorporated in the understanding stage. We further consider emotion in the generation stage by predicting the system conduct in the end-to-end pipeline, as illustrated in Figure 1b. To this end, we build a LLaMA-based end-to-end ToD system that involves emotion in both understanding and generation, with LLaMA-2-7B (Touvron et al., 2023) as the backbone. As illustrated in Figure 1b, it takes dialogue history and the recognised user emotion as input, and then auto-regressively generates the dialogue state, user emotion, system actions, system conduct, and delexicalised natural language response. The response is then lexicalised via database queries based on the intermediately generated dialogue state and system actions. We refer to this end-to-end model as **EmoLLAMA**.

We did not train EmoLLAMA via RL with task and emotion feedback from the user simulator because it would take more than 20 days on an A100 40GB to simulate the same number of dialogues as we did to train the EmoDDPT policy in the modular system. We therefore leave efficient training of LLM-based ToD systems via RL as a future research direction.

4.5 Emotional User Simulation

Traditionally, user simulators interact with the system on the semantic level for efficiency. To capture more fine-grained expressions of system conducts in natural language, we build **langEmoUS** based on EmoUS (Lin et al., 2023). langEmoUS interacts with the system on the natural language level, e.g. it takes the system utterance, user goal, turn information and user persona as inputs and generates user emotion and user utterance. The turn information represents the dialogue progress, i.e. the turn number. Following the setting in Lin et al. (2023), the user persona is extracted from the dialogue history, e.g. if a user is excited to visit a museum in the conversation, then its persona is $\{attraction : excited\}$, when training the user model supervisedly. During inference, the user persona is sampled from the distribution of the corpus.

LangEmoUS achieves macro F1 scores of 0.742 and 0.521 for user sentiment prediction and emotion prediction, respectively, significantly outperforming existing state-of-the-art models (Kim and Lipani, 2022; Lin et al., 2023) (see Appendix B).

5 Experimental Set-up

5.1 Modular System Set-up

EmoLoop This is our proposed modular system with emotion incorporated for understanding, management, and understanding, as outlined in Figure 1a and Figure 2. It includes the following modules: SetSUMBT DST, ContextBERT-ERToD ERC, EmoDDPT policy, and SEC-BART NLG.

EmoDDPT is trained via RL on the natural language level with langEmoUS.

SimpleLoop This is the non-emotion baseline to EmoLoop. It neither predicts user emotion for the state, uses emotion reward to train the policy, nor generates system conduct for emotional response generation. Specifically, it includes the following modules: SetSUMBT DST, DDPT policy, and SC-BART NLG. DDPT is trained via RL on the natural language level with langEmoUS.

5.1.1 Dialogue Policy Optimisation

We implement our system in the ConvLab-3 framework (Zhu et al., 2023). We pre-trained the policy on MultiWOZ 2.1 (Eric et al., 2020), followed by online RL through interaction with our US. During RL, in addition to the emotion reward as outlined in Section 4.2, we set the task reward as -1 in every turn to encourage efficiency, and 80 or -40 for dialogue success or failure. A dialogue is successful if the system provides the requested information to the user and books the correct entities (if possible). For emotional reward, we set $\beta = 2$. We pre-train each policy on MultiWOZ, followed by 15k dialogues with langEmoUS via RL for 6 random seeds. For every 1k dialogues of training, we evaluate the policy for 500 dialogues. We use overall return to select the best checkpoint. All peripheral modules were trained, implemented, and evaluated in the ConvLab-3 environment.

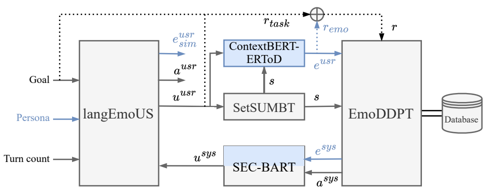


Figure 2: RL training set-up for EmoDDPT.

Language-level RL Training As illustrated in Figure 2, our policy, EmoDDPT, interacts with langEmoUS on the natural-language level where the policy actions and conduct (a^{sys}, e^{sys}) is realised into natural language, u^sys with SEC-BART. The US takes natural-language input and outputs natural-language user utterances u^{usr} after auto-regressively generating the simulation target user emotion u^{usr}_{sim} and user actions a^{usr}. The perceived user emotion e^{usr} and dialogue state s are determined by ContextBERT-ERToD and SetSUMBT respectively.

5.2 End-to-end System Set-up

EmoLLAMA This our proposed end-to-end system as described in Section 4.4.

SimpleLLAMA This is the non-emotional baseline, which is also used in the work of Stricker and Paroubek (2024). Compared with EmoLLAMA, it does not consider user emotions as a part of the model input, nor does it auto-regressively predict user emotion and system conduct.

Both EmoLLAMA and SimpleLLAMA are trained and evaluated with EmoWOZ 2.0 using the environment provided by Stricker and Paroubek (2024) and following default parameters. Their interactive evaluations were set up in the ConvLab-3 environment.

5.3 Evaluation

Corpus Evaluation We report *inform* and *success* rates. Inform rate evaluates if the system provides entities from the database that fulfill user's constraints. Success rate assesses if the system delivers all information requested by the user. To generate each system response, the ground-truth dialogue history was used as system input.

Interactive Evaluation For interactive evaluation, our systems interact with langEmoUS. We report the *success* rate and the average user *sentiment* in simulated dialogues to account for user emotional experience. Specifically, the turn-level sentiment score is $+1$ if the user emotion is positive, 0 if neutral and -1 if negative. User sentiment is determined by the ERC.

Human Trial We set up a human trial using the DialCrowd toolkit (Huynh et al., 2022) on the Amazon Mechanical Turk platform. We set up two pairs of comparison: 1) SimpleLLAMA vs. EmoLLAMA and 2) SimpleLoop vs. EmoLoop. Volunteers are presented with randomly generated single or multi-domain goals. A goal contains a set of constraints for entities that the user should be looking for (e.g. the price range and the location of a restaurant) and specifies the information they should extract from the system (e.g. the phone number and booking reference of the restaurant). Given a goal, volunteers would need to talk to each system to fulfill the goal. They then give ratings to each of them based on objective (whether the goal has been fulfilled) and subjective metrics (how they feel about the system). Survey questions include objective task success and subjective user

System	Type	Corpus		User Simulator		Human	
		Inform	Success	Success	Sentiment	Success	Sentiment Rating
SimpleLLAMA	End-to-end	0.785	0.705	0.330	0.214	0.819	3.97
EmoLLAMA	End-to-end	**0.833**	**0.760**	0.342	**0.250**	**0.894**	**4.16**
SimpleLoop	Modular	0.700	0.621	0.556	0.337	0.798	3.85
EmoLoop	Modular	**0.753**	0.635	0.531	**0.405**	**0.917**	**4.15**

Table 1: System evaluation, including corpus-based evaluation, interaction with user simulator and human trial. Values in bold mean best scores with statistically significant difference $p < 0.05$.

sentiment. Details of the website interface and survey questions can be found in Appendix D. To obtain more reliable ratings, we filtered out dialogues with poor quality, e.g. containing very short user utterances or non-natural language, and with inconsistent ratings, e.g. system A had better rating in all aspects but overall the rater found system B better. Overall, we collected 203 valid ratings for the SimpleLLAMA-EmoLLAMA comparison and 253 for the SimpleLoop-EmoLoop comparison from 40 unique raters.

6 Results and Discussion

6.1 Corpus Evaluation

Although it is not a common practice to evaluate RL-trained modular ToD systems on a corpus, we provide such results for a basic understanding and comparison with end-to-end systems. Our goal is not beating SOTA on task-related metrics, but examining interactive abilities of the system and the role of emotion in it. As shown in Table 1, incorporating emotion significantly improves inform rate of both types of systems and success rate of the end-to-end system.

It is not surprising that modular systems underperform when compared with end-to-end systems. Modular systems are trained via RL, which allows the policy to explore more diverse dialogue trajectories but diverges from what a policy can learn from the corpus only. This reflects the limitation of corpus evaluation in accounting for ToD system performance, as pointed out by Lubis et al. (2022).

6.2 Evaluation with User Simulator

In interactive evaluation, both EmoLoop and EmoLLAMA perform significantly better in terms of average sentiment than their respective non-emotional baseline while maintaining the same level of success rate. For end-to-end models, despite the fact that they are not optimised via RL with the simulated user, the average sentiment in the simulated user also improves significantly.

When comparing performance across system types, modular systems perform better than end-to-end models on task success and simulated user sentiment since modular system policies have been optimised for the simulated user via RL. SimpleLLAMA and EmoLLAMA, trained via SL only, cannot adequately cope with the more diverse user goals and situations of the simulated user. This motivates our future work to leverage the simulated user and to train end-to-end systems via RL.

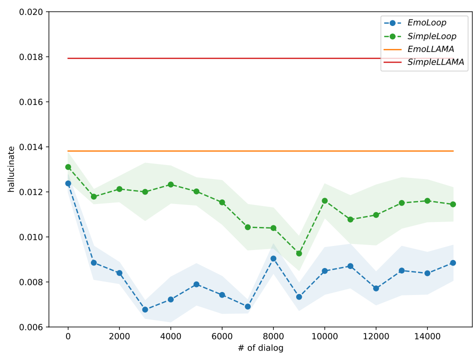


Figure 3: The average hallucination rate of modular systems during RL training with langEmoUS. For end-to-end systems, we report hallucination rate after SL.

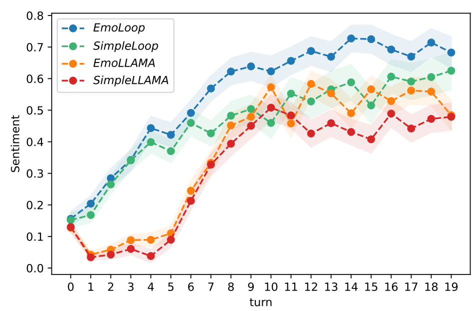


Figure 4: Average sentiment at different turn positions during language-level interaction with langEmoUS.

Hallucination In ToD, a hallucination is defined as a value in the system response that is not supposed to be informed according to system actions. As shown in Figure 3, the hallucination rate of each type of systems is improved as emotion is incorporated into the pipeline. The hallucination rate is lowered from 1.8% for SimpleLLAMA to 1.4% for EmoLLAMA. We observe that end-to-end systems are more prone to the hallucination problem than modular systems as slot placeholders in the delexicalised end-to-end system response do not always match the intermediately generated dialogue actions. Hallucination rates of SimpleLoop and EmoLoop are around 1.3% at the beginning of the interactive RL training and continue to improve as the RL progresses.

Progression of User Sentiment in Dialogues Figure 4 shows the average sentiment of langEmoUS at each turn of interactions with our systems. The sentiment level of langEmoUS becomes more positive as the dialogue progresses and moves towards user goal completion in all systems. The primary difference between modular systems and end-to-end systems is that in earlier turns, modular systems are able to satisfy the simulated user better, as illustrated in higher and more positive sentiment level before turn 8.

6.3 Human Trials

We carried out human trials to compare two pairs of systems in Table 1. Within each pair of comparison, the emotion-incorporating model significantly outperforms its non-emotion version in terms of both the success rate and user sentiment. This further confirms our findings from corpus and user simulator evaluations. Example dialogue excerpts are given in Appendix D.3 to exemplify how emotional ToD systems made use of affective conduct in case of neutral and unsuccessful interactions.

Although human ratings across system types are not directly comparable, it is noteworthy that the absolute improvement from SimpleLLAMA to EmoLLAMA (ΔSuccess $= 0.075$, ΔSentiment $= 0.19$) is smaller than that from SimpleLoop to EmoLoop (ΔSuccess $= 0.119$, ΔSentiment $= 0.30$). Such difference can be attributed to the lack of RL training in LLM-based systems.

6.4 Ablation Study

We ablate our emotional modular and end-to-end systems by incorporating emotion in different parts of the pipeline. Table 2 summarises their interactive performance with langEmoUS.

For both modular systems and end-to-end systems, incorporating emotion does not significantly

System	Und	Gen	Man	Success	Sentiment
SimpleLLAMA	-	-	-	0.330	0.214
	+	-	-	0.360	0.233
	-	+	-	0.373	0.229
EmoLLAMA	+	+	-	0.342	**0.250**
SimpleLoop	-	-	S	0.556	0.337
	+	-	S+E	0.559	0.354
	-	+	S	0.543	0.361
EmoLoop	+	+	S+E	0.531	**0.405**

Table 2: Success and average user sentiment of systems from the interactive evaluation with langEmoUS. +/- means whether emotion is involved in the corresponding ToD stage: **Und**erstanding, **Man**agement, or **Gen**eration. For Management, "-" means the system is trained via SL,"S" and "E" mean training via RL with success reward and emotion reward respectively.

change task success with the user simulator ($p > 0.5$). The average user sentiment does improve slightly as emotion is introduced in understanding (plus management) and generation. Yet, the improvement from the non-emotional base system only becomes significant when emotion is added to all ToD stages. This highlights the importance of considering emotion in the whole ToD loop: it is necessary not only to understand user emotion but also to make use of it for dialogue management and respond with the appropriate conduct. *

Figure 5 illustrates the change in the average sentiment of the simulated user during RL. At the beginning, average sentiments of modular systems fall in the similar range as SL-trained end-to-end systems, and are then further improved by RL. This highlights the importance of task success and emotion feedback signal for RL in ToD systems.

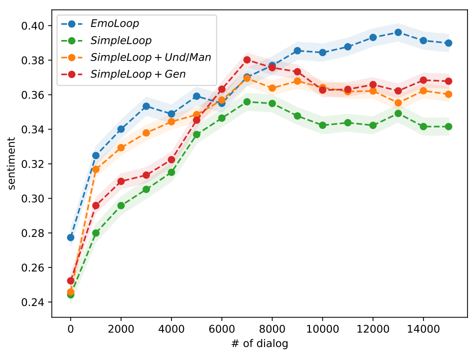


Figure 5: The average sentiment of langEmoUS during RL training of modular policy.

*See Appendix E.1 for ablation study on EmoLoop with SL policy. A similar trend has been observed.

7 Conclusion

In this work, we incorporate emotion into the complete ToD processing loop, involving understanding, management, and generation. To achieve this, we first enrich the EmoWOZ dataset with system conduct labels to construct EmoWOZ 2.0. We then build modular and end-to-end ToD systems, as well as emotional user simulators with the newly collected dataset. We train the modular system policy via RL with the emotional user simulator and the end-to-end system via SL on EmoWOZ 2.0. Through interactive evaluation with both simulated and human users, we show that incorporating emotion into ToD systems can improve user's emotional experience as well as task success.

There is still a long way to go from our work to the perfect emotional ToD system. Yet, we show our method as a promising avenue to achieve this ultimate goal. In our study, we directly translate user emotion labels into valence scores on a linear scale as a reward for RL. We believe that utilising the full set of user emotion labels for diverse reward would be a promising future direction.

We hope that with our work, we can motivate future research efforts to look at user experience beyond task success for ToDs and bring about insights to other task-oriented conversation settings. We would also like to highlight the opportunities in further improving LLM-based end-to-end ToD systems via RL, combining established approaches for policy training in modular systems and recent advancements in LLM research in other applications.

8 Limitations

One of the main limitations of modular ToD systems is the error accumulation in the pipeline for both modular and end-to-end systems. In modular systems, since each module is trained with a dataset associated with a limited ontology, the concepts that the system can understand and express are also limited. Although the DDPT policy, SetSUMBT DST, and many other models such as Trippy-R (Heck et al., 2022) are built with the ability to handle out-of-domain requests, the generalisability and robustness of ToD systems are still challenges in the field that is yet to be solved.

All system modules have been trained in a supervised fashion on EmoWOZ 2.0. Therefore, the dataset contains limited dialogue situations and inherent bias. As seen in the dialogue examples in the appendix, the emotional responses are also limited. Yet, EmoWOZ 2.0 is the best resource we have at the moment. Data augmentation has been applied when training the NLG and the ERC model to mitigate the lack of diversity in the dataset. The RL training of the policy also allows the policy to explore more diverse dialogue trajectories. For the user simulator, considering data augmentation and more attributes of users, e.g. a more fine-grained user persona from chit-chat, would be a potential future direction to improve the diversity in simulated user behaviours.

Although LLMs can have better performance on each ToD modelling task and therefore could potentially serve as more powerful modules in EmoLoop, we did not move in this direction since their high computing resource requirement and slow inference speed would hinder their integration into our systems for interactive training and evaluation. Training modular system policy with langEmoUS for 15k dialogues on one Nvidia GeForce RTX 2080 Ti takes around 40 hours. The training time and memory required will be significantly increased if modular systems use LLM-based modules. On the other hand, while LLM-based end-to-end systems may provide a bypass since one LLM is sufficient, implementing RL training on such systems to further leverage task success and emotion signals from the user simulator is another computationally expensive challenge that are yet to solve.

Some of our generative system modules are based on pre-trained language models. Although we have not been reported any harmful generations in the human trail, there is still the possibility for unexpected behaviour when this system is deployed and tested on a very large scale.

For human evaluation, we conducted experiments on Amazon Mechanical Turk platform rather than deployed our systems in the production environment. The participants, despite coming from different countries, are from covering all demographics.

9 Ethics Statement

Models, codes and datasets were used in accordance with their respective licenses, terms of use and intended use. The data that we used and generated does not contain any information that names or uniquely identifies individual people or offensive content. The model we used for generating augmented samples has implemented training objectives for enhanced safety (Appendix C). Systems we used for interaction with real users were very unlikely to generate offensive content as they were fine-tuned on large-scale training data to convey a limited scope of semantic concepts. No offensive content was reported by human users nor observed in post-hoc inspection.

For system conduct annotation, annotators were required to read and agree with our statement of consent for data use before the task. Annotators were paid fairly according to the local regulations of our research institute. We ensured swift communication with annotators so that their concerns were addressed as soon as possible. For poor-quality annotations, we still pay the annotators for their time but block them from our task to ensure data quality and collection efficiency. All annotations are anonymised.

The data annotation and interactive human trial, which involves decision making based on human emotions, have been approved by the ethics review board of the research institute. The proposed system learns how to manipulate human emotional state. Although the system is trained to elicit positive user emotion, this could still be of potential ethical concern and would require greater deliberation when deployed in real-life and more complex scenario.

10 Acknowledgement

S. Feng, N. Lubis, and M. Heck are supported by funding provided by the Alexander von Humboldt Foundation in the framework of the Sofja Kovalevskaja Award endowed by the Federal Ministry of Education and Research. H-C. Lin and C. van Niekerk are supported by the Ministry of Culture and Science of North Rhine-Westphalia within the framework of the Lamarr Fellow Network. C. Geishauser, B. Ruppik, and R. Vukovic are supported by funds from the European Research Council (ERC) provided under the Horizon 2020 research and innovation programme (Grant agreement No. STG2018804636). Computing resources were provided by Google Cloud.

References

Magda B. Arnold. 1960. *Emotion and personality. Vol. I. Psychological aspects.* Columbia Univer. Press.

Trung Bui, Job Zwiers, Mannes Poel, and Anton Nijholt. 2010. Affective dialogue management using factored pomdps. *Studies in Computational Intelligence*, 281:207–236.

Zhiyu Chen, Bing Liu, Seungwhan Moon, Chinnadhurai Sankar, Paul Crook, and William Yang Wang. 2022. KETOD: Knowledge-enriched task-oriented dialogue. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2581–2593, Seattle, United States. Association for Computational Linguistics.

Mark Davis. 2018. *Empathy: A Social Psychological Approach*. Routledge, New York.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Wieland Eckert, Esther Levin, and Roberto Pieraccini. 1997. User modeling for spoken dialogue system evaluation. In *Proceedings of the Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pages 80–87, Santa Barbara, CA, USA. IEEE.

Paul Ekman. 1992. An argument for basic emotions. *Cognition & Emotion*, 6:169–200.

Mihail Eric, Rahul Goel, Shachi Paul, Abhishek Sethi, Sanchit Agarwal, Shuyang Gao, Adarsh Kumar, Anuj Goyal, Peter Ku, and Dilek Hakkani-Tur. 2020. MultiWOZ 2.1: A consolidated multi-domain dialogue dataset with state corrections and state tracking baselines. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 422–428, Marseille, France. European Language Resources Association.

Shutong Feng, Nurul Lubis, Christian Geishauser, Hsien-chin Lin, Michael Heck, Carel van Niekerk, and Milica Gasic. 2022. EmoWOZ: A large-scale corpus and labelling scheme for emotion recognition in task-oriented dialogue systems. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 4096–4113, Marseille, France. European Language Resources Association.

Shutong Feng, Nurul Lubis, Benjamin Ruppik, Christian Geishauser, Michael Heck, Hsien-chin Lin, Carel van Niekerk, Renato Vukovic, and Milica Gasic. 2023a. From chatter to matter: Addressing critical steps of emotion recognition learning in task-oriented dialogue. In *Proceedings of the 24th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 85–103, Prague, Czechia. Association for Computational Linguistics.

Shutong Feng, Guangzhi Sun, Nurul Lubis, Chao Zhang, and Milica Gašić. 2023b. Affect recognition in conversations using large language models. *CoRR*, abs/2309.12881.

Christian Geishauser, Carel van Niekerk, Hsien-chin Lin, Nurul Lubis, Michael Heck, Shutong Feng, and Milica Gašić. 2022. Dynamic dialogue policy for continual reinforcement learning. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 266–284, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Christian Geishauser, Carel van Niekerk, Nurul Lubis, Hsien-chin Lin, Michael Heck, Shutong Feng, Benjamin Ruppik, Renato Vukovic, and Milica Gašić. 2024. Learning with an open horizon in ever-changing dialogue circumstances. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 32:2352–2366.

Alicia A. Grandey, Lori S. Goldberg, and S. Douglas Pugh. 2011. Why and when do stores with satisfied employees have satisfied customers?: The roles of responsiveness and store busyness. *Journal of Service Research*, 14(4):397–409.

Michael Heck, Nurul Lubis, Carel van Niekerk, Shutong Feng, Christian Geishauser, Hsien-Chin Lin, and Milica Gašić. 2022. Robust dialogue state tracking with weak supervision and sparse data. *Transactions of the Association for Computational Linguistics*, 10:1175–1192.

Ehsan Hosseini-Asl, Bryan McCann, Chien-Sheng Wu, Semih Yavuz, and Richard Socher. 2020. A simple language model for task-oriented dialogue. In *Advances in Neural Information Processing Systems*, volume 33, pages 20179–20191. Curran Associates, Inc.

Jessica Huynh, Ting-Rui Chiang, Jeffrey Bigham, and Maxine Eskenazi. 2022. DialCrowd 2.0: A quality-focused dialog system crowdsourcing toolkit. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 1256–1263, Marseille, France. European Language Resources Association.

Dan Jurafsky and James H. Martin. 2009. *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Pearson Prentice Hall, Upper Saddle River, N.J.

To Eun Kim and Aldo Lipani. 2022. A multi-task based neural model to simulate users in goal oriented dialogue systems. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '22, page 2115–2119, New York, NY, USA. Association for Computing Machinery.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*, San Diego, CA, USA.

Florian Kreyssig, Iñigo Casanueva, Paweł Budzianowski, and Milica Gašić. 2018. Neural user simulation for corpus-based policy optimisation of spoken dialogue systems. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 60–69, Melbourne, Australia. Association for Computational Linguistics.

Wai-Chung Kwan, Hong-Ru Wang, Hui-Min Wang, and Kam-Fai Wong. 2023. A survey on recent advances and challenges in reinforcement learning methods for task-oriented dialogue policy learning. *Machine Intelligence Research*, 20(3):318–334.

Richard S. Lazarus. 1966. *Psychological stress and the coping process.* McGraw-Hill, New York.

Esther Levin and Roberto Pieraccini. 1997. A stochastic model of computer-human interaction for learning dialogue strategies. In *Procceedings of the 5th European Conference on Speech Communication and Technology (Eurospeech 1997)*, pages 1883–1886, Rhodes, Greece.

Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.

Zaijing Li, Ting-En Lin, Yuchuan Wu, Meng Liu, Fengxiao Tang, Ming Zhao, and Yongbin Li. 2023. UniSA: Unified generative framework for sentiment analysis. In *Proceedings of the 31st ACM International Conference on Multimedia*, MM '23, page 6132–6142, New York, NY, USA. Association for Computing Machinery.

Hsien-Chin Lin, Shutong Feng, Christian Geishauser, Nurul Lubis, Carel van Niekerk, Michael Heck, Benjamin Ruppik, Renato Vukovic, and Milica Gašić. 2023. EmoUS: Simulating user emotions in task-oriented dialogues. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '23, page 2526–2531, New York, NY, USA. Association for Computing Machinery.

Hsien-chin Lin, Christian Geishauser, Shutong Feng, Nurul Lubis, Carel van Niekerk, Michael Heck, and Milica Gasic. 2022. GenTUS: Simulating user behaviour and language in task-oriented dialogues with generative transformers. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 270–282, Edinburgh, UK. Association for Computational Linguistics.

Hsien-chin Lin, Nurul Lubis, Songbo Hu, Carel van Niekerk, Christian Geishauser, Michael Heck, Shutong Feng, and Milica Gasic. 2021. Domain-independent user simulation with transformers for task-oriented dialogue systems. In *Proceedings of the 22nd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 445–456, Singapore and Online. Association for Computational Linguistics.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. RoBERTa: A robustly optimized BERT pretraining approach. *CoRR*, abs/1907.11692v1.

Nurul Lubis, Christian Geishauser, Hsien-chin Lin, Carel van Niekerk, Michael Heck, Shutong Feng, and Milica Gasic. 2022. Dialogue evaluation with offline reinforcement learning. In *Proceedings of the 23rd Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 478–489, Edinburgh, UK. Association for Computational Linguistics.

John Mendonca, Patrícia Pereira, Miguel Menezes, Vera Cabarrão, Ana C Farinha, Helena Moniz, Alon Lavie, and Isabel Trancoso. 2023. Dialogue quality and emotion annotations for customer support conversations. In *Proceedings of the Third Workshop on Natural Language Generation, Evaluation, and Metrics (GEM)*, pages 9–21, Singapore. Association for Computational Linguistics.

Andrew Ortony, Gerald L. Clore, and Allan Collins. 1988. *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge.

Baolin Peng, Chenguang Zhu, Chunyuan Li, Xiujun Li, Jinchao Li, Michael Zeng, and Jianfeng Gao. 2020. Few-shot natural language generation for task-oriented dialog. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 172–182, Online. Association for Computational Linguistics.

Rosalind W. Picard. 2000. *Affective Computing*. The MIT Press, Cambridge, Mass.

Fuji Ren, Yu Wang, and Changqin Quan. 2015. TFSM-based dialogue management model framework for affective dialogue systems. *IEEJ Transactions on Electrical and Electronic Engineering*, 10(4):404–410.

Armand Stricker and Patrick Paroubek. 2024. A Unified Approach to Emotion Detection and Task-Oriented Dialogue Modeling. In *IWSDS*, Sapporo (Japon), Japan.

Kai Sun, Seungwhan Moon, Paul Crook, Stephen Roller, Becka Silvert, Bing Liu, Zhiguang Wang, Honglei Liu, Eunjoon Cho, and Claire Cardie. 2021. Adding chit-chat to enhance task-oriented dialogues. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1570–1583, Online. Association for Computational Linguistics.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton-Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurélien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. 2023. Llama 2: Open foundation and fine-tuned chat models. *CoRR*, abs/2307.09288.

Carel van Niekerk, Andrey Malinin, Christian Geishauser, Michael Heck, Hsien-chin Lin, Nurul Lubis, Shutong Feng, and Milica Gasic. 2021. Uncertainty measures in neural belief tracking and the effects on dialogue policy performance. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 7901–7914, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Tom Young, Frank Xing, Vlad Pandelea, Jinjie Ni, and E. Cambria. 2021. Fusing task-oriented and open-domain dialogues in conversational agents. In *AAAI Conference on Artificial Intelligence*. Association for the Advancement of Artificial Intelligence.

Rui Zhang, Zhenyu Wang, Mengdan Zheng, Yangyang Zhao, and Zhenhua Huang. 2021. Emotion-sensitive deep dyna-q learning for task-completion dialogue policy learning. *Neurocomputing*, 459:122–130.

Hui Zhu, Xv Wang, Zhenyu Wang, and Kai Xv. 2024. ESDP: An emotion-sensitive dialogue policy for task-oriented dialogue system.

Qi Zhu, Christian Geishauser, Hsien-chin Lin, Carel van Niekerk, Baolin Peng, Zheng Zhang, Shutong Feng, Michael Heck, Nurul Lubis, Dazhen Wan, Xiaochen Zhu, Jianfeng Gao, Milica Gasic, and Minlie Huang. 2023. ConvLab-3: A flexible dialogue system toolkit based on a unified data format. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 106–123, Singapore. Association for Computational Linguistics.

A EmoWOZ 2.0 Construction

A.1 Annotation Scheme Justification

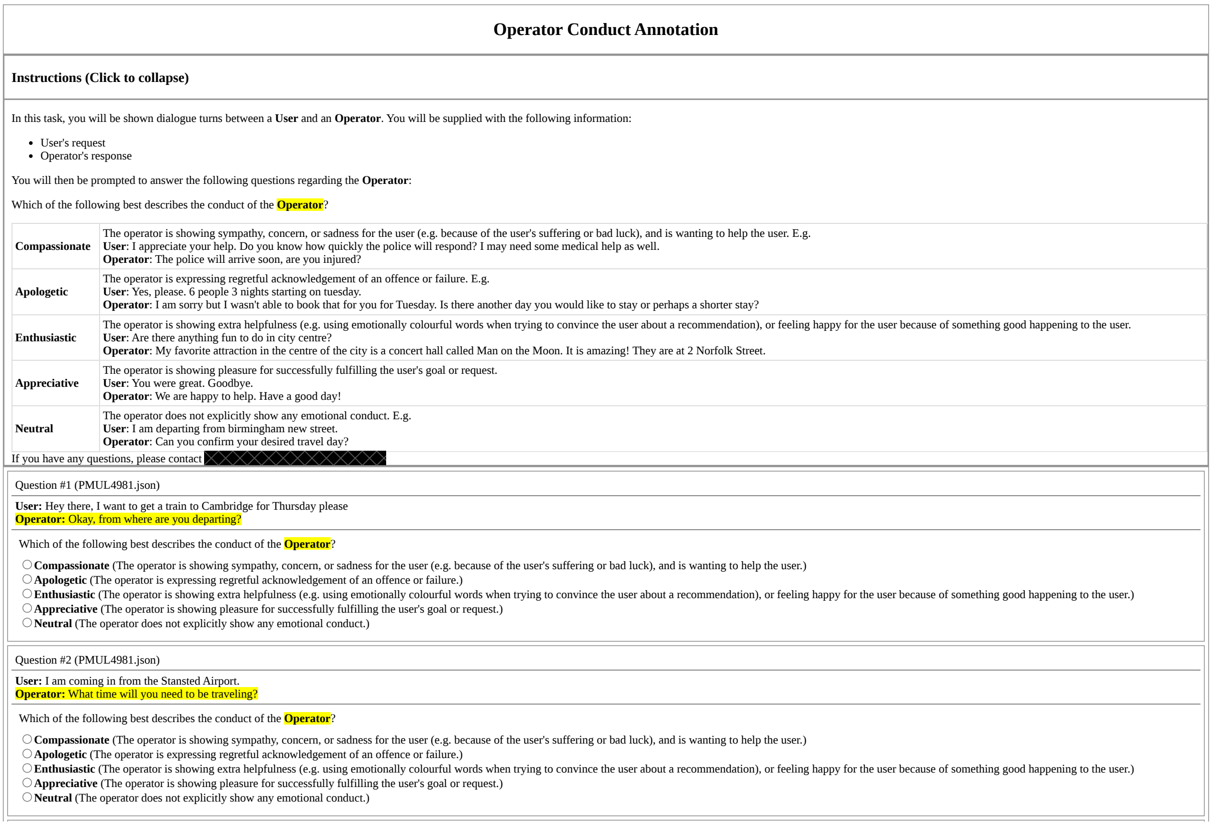
Under the framework of the OCC emotion model and the definition of emotional empathy that the observer shares the emotional state of another person (Davis, 2018), we can derive the corresponding emotional response from the system. Considering the following user emotion and situation where:

The user is labelled as *Fearful*, or feeling negative because of an event which has negative consequences on the user his or herself (as defined in EmoWOZ).

An empathetic operator would share the same feeling as the user (therefore also feeling negative). Yet, the feeling in the operator is elicited by an event which has negative consequences on the user (the other party). This feeling is defined as pity, or compassionate in the OCC model.

A.2 Annotation Interface

We adopted the same annotation set-up, annotator selection criteria, and quality assurance approaches as outlined by Feng et al. (2022). Each utterance is annotated by three annotators, who were provided with the entire preceding dialogue history when annotating the current utterance. Annotators were English speakers. The final label was obtained from majority voting. When the agreement could not be reached, a fourth annotator was introduced. Overall, 54 crowd workers have contributed to our study.



Operator Conduct Annotation

Instructions (Click to collapse)

In this task, you will be shown dialogue turns between a **User** and an **Operator**. You will be supplied with the following information:

- User's request
- Operator's response

You will then be prompted to answer the following questions regarding the **Operator**:

Which of the following best describes the conduct of the **Operator**?

Compassionate	The operator is showing sympathy, concern, or sadness for the user (e.g. because of the user's suffering or bad luck), and is wanting to help the user. E.g. **User**: I appreciate your help. Do you know how quickly the police will respond? I may need some medical help as well. **Operator**: The police will arrive soon, are you injured?
Apologetic	The operator is expressing regretful acknowledgement of an offence or failure. E.g. **User**: Yes, please. 6 people 3 nights starting on tuesday. **Operator**: I am sorry but I wasn't able to book that for you for Tuesday. Is there another day you would like to stay or perhaps a shorter stay?
Enthusiastic	The operator is showing extra helpfulness (e.g. using emotionally colourful words when trying to convince the user about a recommendation), or feeling happy for the user because of something good happening to the user. **User**: Are there anything fun to do in city centre? **Operator**: My favorite attraction in the centre of the city is a concert hall called Man on the Moon. It is amazing! They are at 2 Norfolk Street.
Appreciative	The operator is showing pleasure for successfully fulfilling the user's goal or request. **User**: You were great. Goodbye. **Operator**: We are happy to help. Have a good day!
Neutral	The operator does not explicitly show any emotional conduct. E.g. **User**: I am departing from birmingham new street. **Operator**: Can you confirm your desired travel day?

If you have any questions, please contact [redacted]

Question #1 (PMUL4981.json)

User: Hey there, I want to get a train to Cambridge for Thursday please
Operator: Okay, from where are you departing?

Which of the following best describes the conduct of the **Operator**?

- ○ **Compassionate** (The operator is showing sympathy, concern, or sadness for the user (e.g. because of the user's suffering or bad luck), and is wanting to help the user.)
- ○ **Apologetic** (The operator is expressing regretful acknowledgement of an offence or failure.)
- ○ **Enthusiastic** (The operator is showing extra helpfulness (e.g. using emotionally colourful words when trying to convince the user about a recommendation), or feeling happy for the user because of something good happening to the user.)
- ○ **Appreciative** (The operator is showing pleasure for successfully fulfilling the user's goal or request.)
- ○ **Neutral** (The operator does not explicitly show any emotional conduct.)

Question #2 (PMUL4981.json)

User: I am coming in from the Stansted Airport.
Operator: What time will you need to be traveling?

Which of the following best describes the conduct of the **Operator**?

- ○ **Compassionate** (The operator is showing sympathy, concern, or sadness for the user (e.g. because of the user's suffering or bad luck), and is wanting to help the user.)
- ○ **Apologetic** (The operator is expressing regretful acknowledgement of an offence or failure.)
- ○ **Enthusiastic** (The operator is showing extra helpfulness (e.g. using emotionally colourful words when trying to convince the user about a recommendation), or feeling happy for the user because of something good happening to the user.)
- ○ **Appreciative** (The operator is showing pleasure for successfully fulfilling the user's goal or request.)
- ○ **Neutral** (The operator does not explicitly show any emotional conduct.)

Figure A.1: Web-interface for conduct annotation.

A.3 Annotator Confusion Matrix

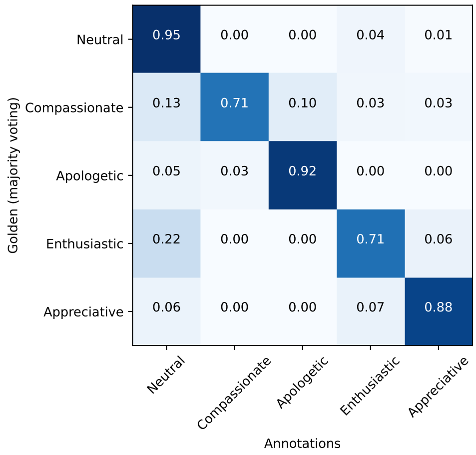


Figure A.2: Annotator confusion matrix.

A.4 System Conduct Distribution

Conduct	**Count**	**Proportion**
Neutral	52,236	73.0%
Appreciative	9,763	13.6%
Enthusiastic	6,364	8.9%
Apologetic	3,049	4.3%
Compassionate	112	0.2%

Table A1: Conduct distribution in MultiWOZ.

B User Simulator Implementation Details

Following the setting in Lin et al. (2023), the input and output of langEmoUS are represented as JSON-formatted strings, which are composed of tokens in natural language. We initialised our model based on the BART model (Lewis et al., 2020) and fine-tuned it on our EmoWOZ 2.0 dataset. We optimised our model with Adam (Kingma and Ba, 2015), where the learning rate is $2e^{-5}$ for 5 epochs. As shown in Table B1, langEmoUS achieves state-of-the-art performance on user sentiment and emotion prediction.

Model	**Sentiment**	**Emotion**
SatActUtt (Kim and Lipani, 2022)	0.379	-
EmoUS (Lin et al., 2023)	0.693	0.501
langEmoUS	**0.742**	**0.521**

Table B1: Performance for emotion and sentiment prediction of different models by measuring macro-F1 score.

C Natural Language Generator Implementation Details

C.1 NLG Training

C.1.1 Training Configuration

We trained SC-BART and SEC-BART on EmoWOZ 2.0. We trained our model with Adam optimiser for standard cross entropy loss where the learning rate was set to $2e^{-5}$ for 5 epochs (with an early-stopping criterion based on the loss in the validation set) and a batch size of 16. During inference, we set the temperature to 0.9 and a beam number of 2 to promote some degree of diversity.

C.1.2 Prompt Template

Our NLG models take the following input: previous user utterance u_t, dialogue semantic actions a_t, and conduct e_t^{sys} (for SEC-BART only). The prompt template is shown as follows:

SEC-BART Given the previous user request "$\{u_t\}$", the natural language realisation of dialogue action "$\{a_t\}$" with a/an "$\{e_t^{sys}\}$" conduct is

SC-BART Given the previous user request "$\{u_t\}$", the natural language realisation of dialogue action "$\{a_t\}$" is

Given the prompt, the model predicted the probability distribution for a sequence of tokens. The output target is the corresponding ground-truth system response in EmoWOZ 2.0.

C.1.3 Model Performance

Model	BLEU ↑	SER ↓
SC-GPT (Peng et al., 2020)	33.6	4.8
T5NLG (Zhu et al., 2023)	35.8	3.7
SC-BART	**35.9**	3.9
SEC-BART	34.9	**3.6**

Table C1: NLG Performance.

C.2 Data Augmentation

C.2.1 Augmented Sample Collection

Since the conduct distribution in EmoWOZ 2.0 is heavily imbalanced, we leveraged large language models for data augmentation. We selected system utterances with neutral conduct as the source to paraphrase for a target non-neutral conduct. We used LLaMA-2-13b-chat model (Touvron et al., 2023). We used the following prompt:

Given the user request "$\{u_t^{usr}\}$" and the operator response action "$\{a_t\}$", please paraphrase the operator response "$\{u_{t,groundtruth}^{sys}\}$" in a more "$\{e_{t,target}^{sys}\}$" way? Please only give the answer, in less than $2 \times len(u_{t,groundtruth}^{sys})$ tokens and enclosed with [RESP][/RESP].

We also experimented with ICL but the model tends to over-fit on the ICL samples. We therefore let it paraphrase in an zero-shot set-up to best explore its knowledge from pre-training for better diversity in the expression.

C.2.2 Augmented Sample Selection

Since the model does not always follow the target conduct. For example, the large language model (LLM) would find some action-conduct combinations unreasonable. We therefore applied filtering on the LLM-generated samples.

Conduct Expressiveness We trained an ensemble of 10 ContextBERT-ERToD models for conduct classification on EmoWOZ 2.0. The classifier reports an average weighted F1 score of 81.8% without neutral. We then used majority voting from the classifier ensemble to correct the original target conduct when generating the sample.

Faithfulness to Semantic Action We used the rule-based script in ConvLab-3 to evaluate NLG slot error rates in the paraphrased output based on the dialogue actions in the prompt. If there are slot errors in the output, we drop the sample.

Overall, we obtained 949 samples for *Compassionate*, 900 for *Apologetic*, 2274 for *Enthusiastic*, and 490 for *Appreciative*.

D Human Evaluation

D.1 Web Interface

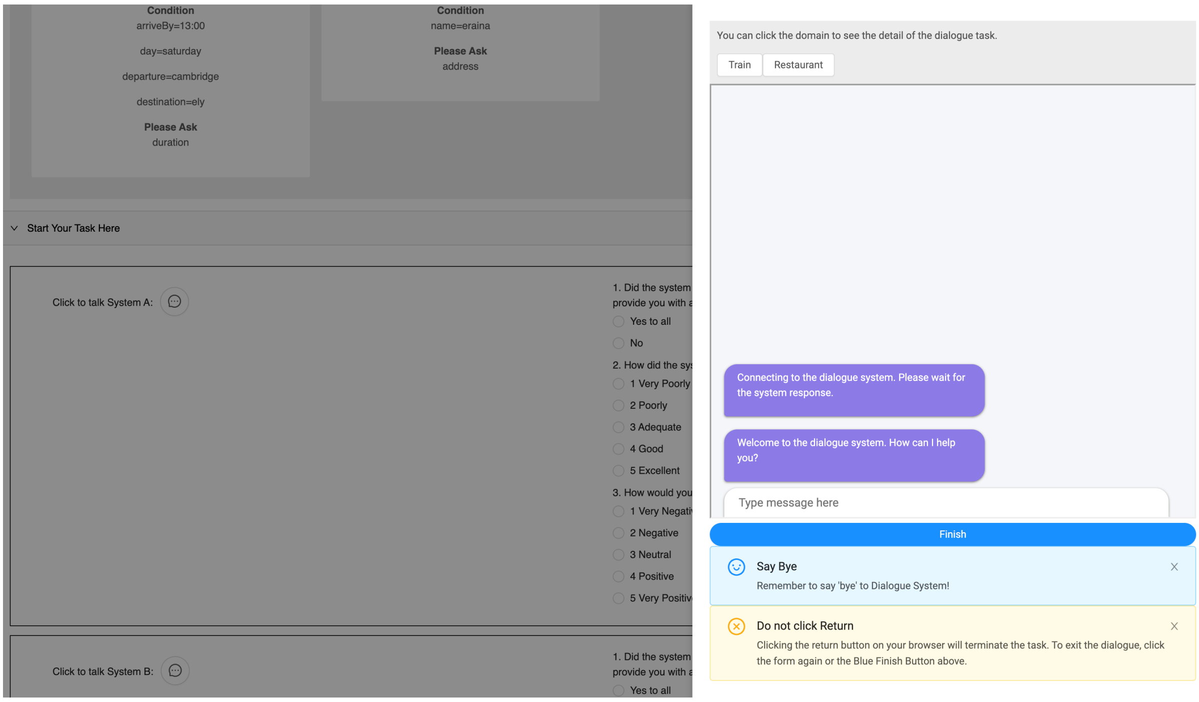


Figure D.1: The web interface for human trial.

D.2 Survey Questions

D.2.1 Question 1 - Task Success

Question Did the system find what you look for? Did it provide all the information that you need? If you ask for a booking, did it provide you with a reference number?

Multiple Choices (A) Yes to all; (B) No.

D.2.2 Question 2 - Sentiment Rating

Question How would you rate your sentiment after the conversation?

Multiple Choices (A) Very Negative; (B) Negative; (C) Neutral; (D) Positive; (E) Very Positive.

D.3 Dialogue Excerpts

System	Role	Text
EmoLoop	User	Hello, I am looking for an attraction called byard art. I am really excited about my trip to Cambridge!!
	System	I'm sorry, there is byard art in the south. The address is 14 kings parade.
	User	Why are you apologising?
	System	I apologize for the confusion. If you need to contact them, please call 01223464646
SimpleLoop	User	No, I just want to book a hotel for 3 people and 3 nights on friday and I need the address.
	System	I can do that. What is the name of the hotel?
	User	Can you suggest one?
	System	I have booked a taxi for you. It is a toyota and the contact number is 29929243243
	User	Sorry, you are not helping at all
	System	I was able to book that for you. The reference number is 63484139414. What is your destination?

Table D1: Human trial dialogue excerpts from EmoLoop and SimpleLoop in a similar situation where both systems were not performing adequately at the beginning. EmoLoop responded with an improper conduct but realised the mistake and continued to provide information. SimpleLoop did not show any sign of realising the mistake.

System	Role	Text
EmoLoop	User	I need the phone number of the hotel with name ashley hotel.
	System	Sure thing! It is located in the north and the postcode is cb41er and the phone number is 01223350059.
	User	Thanks a lot. I also need a spanish restaurant with moderate price.
	System	I would recommend la tasca. It is located at 14-16 bridge street.
SimpleLoop	User	Alright, I am also interested in a museum in the city centre.
	System	I recommend broughton house gallery. It is a museum. Their phone number is 01223314960.
	User	Nice, could you also give me the address and postcode?
	System	The postcode is cb11ln and the phone number is 01223314960.

Table D2: Human trial dialogue excerpts from EmoLoop and SimpleLoop in a similar situation where the user expressed excitement and appreciation. Both systems performed adequately in terms of completing the user's goal. EmoLoop responded in an enthusiastic way whereas SimpleLoop did not respond with affective conduct.

System	Role	Text
EmoLoop	User	I need some information.
	System	What kind of information do you need?
	User	Which department there?
	System	I am sorry, I don't understand. What do you mean?
	User	I need information about haematology.
	System	The Addenbrookes Hospital is located at Hills Rd, Cambridge , postcode cb20qq. The phone number is 01223216297.
SimpleLoop	User	I need information about haematology day unit.
	System	The pizza hut cherry hilton Hospital is located at Hills Rd, Cambridge , cb20qq. The phone number is 01223216297.
	User	Department details?
	System	There are 10 trains that go to london kings cross on sunday. The TR7423 leaves at 05:00 and arrives at 05:51. Would you like me to book that for 6 people?

Table D3: Human trial dialogue excerpts from EmoLLAMA and SimpleLLAMA in a similar situation where the user asked for hospital information. EmoLLAMA, although not understanding the user's vague request, attempted to apologise and clarify the request with the user. The task was successfully completed. SimpleLLAMA carried on with errors and fails the task. (The wrong values, "pizza hut" and "cherry hilton", in the SimpleLLAMA response were due to wrong domains in the dialogue action prediction, which led to corresponding name slots in the response. These name slots were filled during lexicalisation based on actions, resulting in an obviously irrational output.)

System	Role	Text
EmoLoop	User	Recommend a taxi which arrives by 4:15 pm from the fez club to tandoori place.
	System	I'd be happy to help with your request, first I'll need to know where you'll be leaving from, where you're going, and what time you'd like to arrive.
	User	I will be leaving from fez club and going to tandoori palace arriving by 4:15 pm.
	System	Booking was successful. The car type is bmw and the contact number is 19287599612.
SimpleLoop	User	I want to book a taxi for 16:15 from cherry hinton water play to ely train station.
	System	What is your destination?
	User	Ely train station.
	System	Booking completed. Booked car type: ford, Contact number: 46563428252.

Table D4: Human trial dialogue excerpts from EmoLLAMA and SimpleLLAMA in a similar situation where both systems failed to capture all information provided in the user request. EmoLLAMA at first missed the information provided by the user but replied in a compassionate way. The user repeated and then the system provides the correct information. Likewise, SimpleLLAMA missed the destination in the first turn. After the user repeated, the system completed the task for the user. Yet, there is no affective interaction between the user and SimpleLLAMA.

E Further Analysis

E.1 Ablation Study for EmoLoop with Supervised Training Only

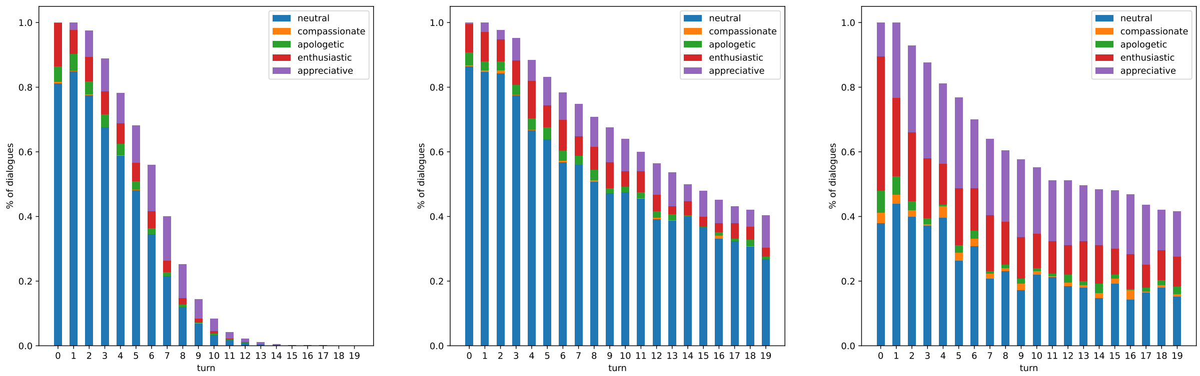
System	Und	Gen	Man	**Success**	**Sentiment**
SimpleLoop-SL	-	-	-	0.512	0.244
	+	-	-	0.494	0.246
	-	+	-	0.493	0.249
EmoLoop-SL	+	+	-	0.516	0.273

Table E1: Success and average user sentiment of our system variants from the interactive evaluation with langEmoUS. +/- means whether the emotion is involved in the corresponding ToD stage: **Und**erstanding, **Man**agement, or **Gen**eration. All systems are trained via SL.

E.2 Impact of Training Set-ups on System Conduct

We investigate how the EmoLoop's affective behaviour is shaped in different stages of training. Figure E.1 shows the distribution of system conduct at different dialogue turns in EmoWOZ 2.0, and policy output during interaction with langEmoUS after supervised pre-training and language-level RL. Comparing Figure E.1a and Figure E.1b suggests that the policy imitates the affective behaviour of operators in the corpus.

After RL, the policy is more inclined to express *enthusiastic* and *appreciative* while expressing *compassionate* and *apologetic* less frequently. This illustrates the affective strategy of the policy to elicit more positive emotions in the simulated user.



(a) Dataset Distribution (b) Supervised Pre-training (c) RL on Language Level

Figure E.1: Distributions of system conduct for different turn positions at different stages of policy training.