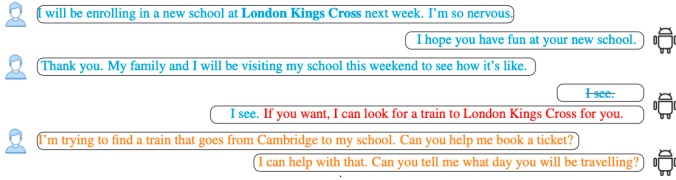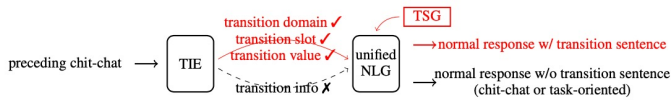## Motivation

1. Dialogue scenarios starting from chit-chat but eventually switching to task-oriented.
2. Detect the implicit user intention of using some task-related services and proactively bridge the connection from CC to TO through generating a transition sentence (red).
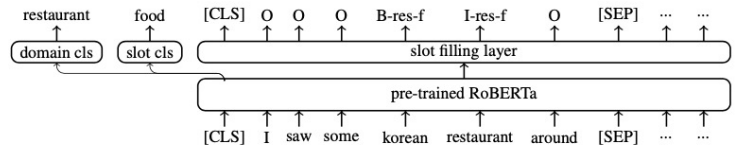3. Benefical for commercial SDSs to proactively offer or sell their task-related services.

Overall Architecture
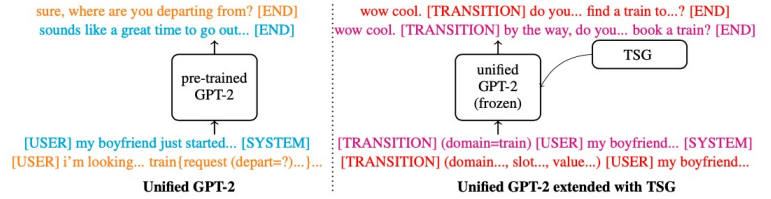

Transition Sentence Augmentation

1. Augment 215 Prepended FusedChat dialogues with a domain guided and a domain-slot-value guided transition sentence respectively.
2. Collect transition sentence templates for different domains and different domain-slot pairs respectively.
3. Randomly asssign templates to the remaining unannotated Prepended FusedChat dialogues.

Transition Info Extractor (TIE)

1. Adapted from the pre-trained RoBERTa, TIE model is built as a NLU model to predict transition domain, transition slot and extract transition value.
2. Jointly modelling slot filling sequence using conditional random fields (CRFs) highly improves the performance of TIE model.

Transition Sentence Generator (TSG)

1. By fine-tuning the pre-trained GPT-2, we leverage the entire FuseChat dataset to train the unified GPT-2 that can reply to both CC and TO user requests.
2. Through leveraging Adapter tuning and Prompt tuning, TS Adapter layers are integrated into the unified GPT-2 and extracted transition information are converted into Prompt tokens to enable the transition sentence generation.
3. Two different transition Prompts are utilized, one only including transition domian, the other including transition domain/slot/value.

Performance Comparison of TIE models

	domain classifier		slot classifier		slot filling		
	accuracy	weighted f1	accuracy	weighted f1	sen_sf_acc	sf_f1	semantic_acc
RoBERTa w/o slot filling	78.57%	79.57%	66.52%	66.84%	–	–	–
joint RoBERTa	82.41%	82.92%	71.86%	73.84%	68.02%	48.64%	61.94%
joint RoBERTa + CRF	**93.71%**	**94.15%**	**82.41%**	**82.30%**	**80.28%**	**61.82%**	**73.67%**

Performance Comparison of Generation models

		Chit-Chat		Task-Oriented			domain TS			domain-slot-value TS		
		Dis-1	Dis-2	BLUE	Meteor	BERTScore (F1)	BERTScore (F1)	transition accuracy	d accuracy	BERTScore (F1)	transition accuracy	d-v accuracy
unified GPT-2		**1.74%**	**12.70%**	**34.77%**	**55.65%**	**93.20%**	–	–	–	–	–	–
retrain	w/o TP	1.67%	11.41%	32.86%	53.52%	92.91%	88.82%	98.25%	58.19%	89.29%	98.97%	30.15%
	w/ TP	1.60%	11.18%	32.58%	53.33%	92.94%	90.19%	98.43%	99.21%	91.70%	98.79%	92.63%
TSG (Houlsby)	w/o TP	**1.74%**	**12.70%**	**34.77%**	**55.65%**	**93.20%**	89.04%	98.67%	62.48%	89.40%	**99.34%**	27.19%
	w/ TP	**1.74%**	**12.70%**	**34.77%**	**55.65%**	**93.20%**	90.28%	**99.40%**	99.15%	**91.84%**	99.21%	**96.80%**
TSG (Pfeiffer)	w/o TP	**1.74%**	**12.70%**	**34.77%**	**55.65%**	**93.20%**	88.90%	97.82%	59.52%	89.33%	98.25%	25.98%
	w/ TP	**1.74%**	**12.70%**	**34.77%**	**55.65%**	**93.20%**	**90.34%**	98.13%	**99.70%**	91.83%	98.43%	96.62%

System-Initiated Transitions from Chit-Chat to Task-Oriented Dialogues with TIE and TSG

Ye Liu, Stefan Ultes, Wolfgang Minker, Wolfgang Maier

Mercedes-Benz · UNIVERSITÄT ULM · universität uulm