	Attraction			Restaurant			Train			Hotel		
Model	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU
MAML	45.5	22.5	9.5	46.0	10.8	7.0	76.3	49.0	**5.7**	48.6	17.7	6.6
DAST	**54.7**	**32.5**	**10.2**	**51.2**	**17.5**	**8.0**	**76.9**	**50.0**	5.5	**49.1**	**25.1**	**7.6**

	Taxi			Police			Hospital			Average		
Model	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU
MAML	-	59.8	7.7	-	41.2	7.0	-	47.6	9.8	54.1	35.5	7.6
DAST	-	**61.8**	7.7	-	**47.2**	**8.2**	-	**48.2**	**10.4**	**58.0**	**40.3**	**8.2**

Table 1: The performance in the metrics of Inform rate, Success rate, and BLEU score on all seven domains from MultiWOZ, as well as the average values over domains. We do not report the Inform rate in domain "Taxi", "Police" and "Hospital" because each of these three domains contains a default task entity. DAST outperforms the MAML baseline in terms of Inform rate and Success rate in every domain and achieves better average BLEU score.

the MultiWOZ dataset. We do not report the Inform rate of "Taxi", "Police" and "Hospital" domain because these three domains have default informable entities, which means the Inform rate is always 100%.

	Slot F1		Act F1	
Domains	MAML	DAST	MAML	DAST
Attraction	31.5	**38.4**	40.0	**41.4**
Restaurant	36.1	**42.5**	32.7	**36.9**
Train	37.5	**41.8**	29.0	**31.0**
Hotel	22.4	**26.0**	27.2	**29.1**
Taxi	-	-	44.0	**45.2**
Police	-	-	51.5	**53.9**
Hospital	-	-	**52.1**	51.1
Average	31.9	**37.2**	39.5	**41.2**

Table 2: The evaluation results on the metrics of Slot F1 and Act F1 in all domains. The DAST outperforms the MAML baseline in terms of Slot F1 in every domain and predicts better dialog acts on average.

The results in the Table 1 show that, for each domain, our model outperforms the baselines in terms of both Inform rate and Success rate. This suggests that the weights generated by the meta-teacher model are beneficial for the student model to optimize adaptation process and achieve better performance in dialog task completion. For the other metric, the BLEU score, our model does not consistently outperform the baselines. This is because the original unweighted loss treats every token in the same way in order to instruct the student model to learn the probabilistic language model. Therefore, the weights from meta-teacher slightly disturb this learning process and consequently reduce the BLEU score. However, our model still achieves better BLEU score than MAML baseline on average, indicating that although slightly affecting learning the language model, the meta-teacher helps the student model to learn new features of the new domain in general.

The performance on the other two metrics is shown in

Model	Slot F1	Act F1	Inform	Success	BLEU
MAML	22.5	50.3	69.2	47.9	12.0
DAST	**23.2**	**59.5**	**89.6**	**61.6**	**12.1**

Table 3: The average performance of all target domains in Schema-Guided Dataset. The DAST outperforms the MAML baseline in terms of all reported metrics

Table 2. Since each of "Taxi", "Police" and "Hospital" domain has a default task entity, which means the explicit state tracking is not required to accomplish the task, we do not report the results on Slot F1 in these three domains. On average, our model outperforms the MAML baseline in both Slot F1 and Act F1, suggesting that the meta-teacher also encourages the dialog model to generate the correct dialog states, which is necessary for database search. And only after searching the database with correct constraints, the dialog system can provide user with a correct task entity.

We also explore the relationship between model performance and the amount of adaptation data. To keep the same setting as SOLOIST (?), we choose only four domains, "Attraction", "Restaurant", "Train" and "Hotel" for evaluation. We randomly sample either 80, 400, 800, or 1600 dialogs in total over four domains. We show all model's results in Table 4. We follow SOLOIST and only report Inform rate, Success rate, and BLEU score. We can find that our method consistently outperforms the baselines, both SOLOIST and MAML, in terms of Inform rate and Success rate. Since these two metrics are closely related to task completion, we believe meta-teacher guides the student model to adapt to new domains more efficiently. On the other hand, the SOLOIST model performs the best in the BLEU score. One main reason is that SOLOIST is fine-tuned based on a pre-trained language model, GPT-2. The pre-trained model makes SOLOIST generates more fluent sentences. However, our teacher-student architecture is generalizable to different student model structures, and can be built based on pre-trained student model as well. In addition, we find that the gap between SOLOIST and our methods reduces with an increasing amount of adaptation data. This is because, with

Model	80			400			800			1600		
	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU	Inform	Success	BLEU
SOLOIST	58.4	35.3	**10.58**	69.3	52.3	**11.8**	69.9	51.9	**14.6**	74	60.1	**15.24**
MAML	62.09	38.36	9.96	72.31	52.91	10.87	74.78	57.71	11.29	76.73	60.61	11.99
DAST	**62.70**	**38.68**	9.49	**74.52**	**54.45**	11.08	**75.92**	**57.72**	11.52	**77.95**	**60.87**	11.97

Table 4: The average performance over four domains, "Attraction", "Restaurant", "Train" and"Hotel", with 80/400/800/1600 dialogs for adaptation. DAST consistently achieves the best performance in the metrics of Inform rate and Success rate, with increasing the amount of adaptation data

enough data, the student model can learn the new domain well without the meta-teacher's guidance. Therefore, the influence of the meta-teacher declines as the number of adaptation dialogs increases.

Table 3 describes the average performance in all the target domains from the Schema-Guided dataset. Our method achieves better performance compared to the MAML baseline in all metrics, suggesting that our method can generalize to different multi-domain dialog datasets.

Case Study and Visualization

Figure 3 lists four example sentences in the restaurant domain from the MultiWOZ dataset, along with their weights assigned by the meta-teacher model. To visualize the weight of each token, we color each token according to its corresponding weight. The larger the weight is, the darker the color is. Since we multiply the weights with token losses to update the student model, the absolute value of the weight can be considered as part of the learning rate. Therefore, we mainly focus on the relative values of weights within the same sentence. And the color intensity only suggests the relative value of weights in the same sentence.

The first two sentences show that our meta-teacher model focuses more on the domain-related tokens like "area" and delexicalized slots such as "[value_area]". One possible reason is that general tokens (like "there are" in the second sentence) have already been learned by the student model in source domains while tokens like "[value_area]" appear less frequently in the source domains. Since large weight amplifies the feedback of the token loss during back-propagation, larger weights for the domain-related tokens encourage the student model to focus on those tokens and quickly learn features of the new domain, leading to more efficient adaptation. In the third case, we find that the delexicalized slots like "[value_range]" and "[value_address]" attracts more attention than domain-related token "address". This is because the domain-related tokens are still possible to be found in other domains. For example, "address" exists in five domains in the MultiWOZ dataset. The last sentence does not contain any domain-specific tokens. Hence, the token weights are close to each other. In this case, the weights do not make much impacts on updating model parameters.

Conclusion and Future work

We propose a domain adaptation method for low-resource task-oriented dialog systems, which incorporates a student-teacher architecture under the meta-learning setting. We present a transformer-based meta-teacher model, which learns to distinguish important tokens under different contexts across source domains during training. As for adaptation, the meta-teacher instructs the student dialog model to pay more attention to influential tokens by assigning weights to token losses, which improves the student model's adaptation performance. We evaluate our method on two popular human-human multi-domain datasets. The results demonstrate that our method reaches state-of-the-art performance in most task-related metrics, compared with MAML and SOLOIST. Since the meta-teacher is built to assign weights to a sequence of generated tokens, our method can be applied to other NLP tasks, such as machine translation and summarization. Furthermore, our meta-teacher model is compatible with other domain adaptation methods, such as MAML and pre-trained models.

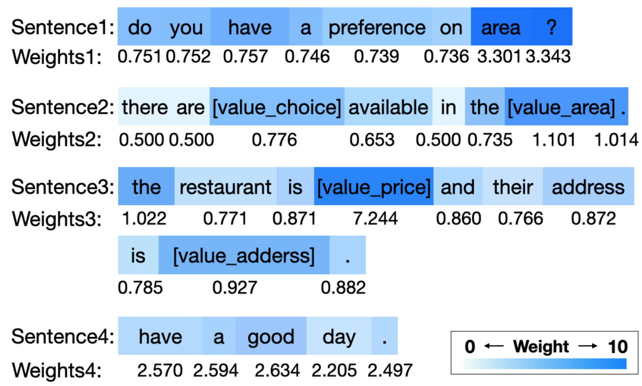

Figure 3: Visualizing weights corresponding to different tokens with different color intensities. The darker the color is, the larger the corresponding token's weight is.

In the future, we aim to extend our method in several directions. First, we plan to include the Success rate and Inform rate into the loss function of the meta-teacher model in a reinforcement learning setting. We believe directly optimize task success metric may lead to better performance. Another direction is to combine the meta-teacher model and pre-trained models to explore the compatibility, as well as replacing GRU-based student model with pre-trained

Cupiditate placeat eveniet deleniti est suscipit vel sequi officiis molestiae a, ipsa consectetur