Towards Low-Resource Semi-Supervised Dialogue Generation with Meta-Learning

Yi Huang, Junlan Feng, Shuo Ma, Xiaoyu Du, Xiaoting Wu
JIUTIAN Team, China Mobile Research
{huangyi, fengjunlan}@chinamobile.com,
{mashuo, duxiaoyu, wuxiaoting}@chinamobile.com

Abstract

In this paper, we propose a meta-learning based semi-supervised explicit dialogue state tracker (SEDST) for neural dialogue generation, denoted as MEDST. Our main motivation is to further bridge the chasm between the need for high accuracy dialogue state tracker and the common reality that only scarce annotated data is available for most real-life dialogue tasks. Specifically, MEDST has two core steps: meta-training with adequate unlabelled data in an automatic way and meta-testing with a few annotated data by supervised learning. In particular, we enhance SEDST via entropy regularization, and investigate semi-supervised learning frameworks based on model-agnostic meta-learning (MAML) that are able to reduce the amount of required intermediate state labelling. We find that by leveraging un-annotated data in meta-way instead, the amount of dialogue state annotations can be reduced below 10% while maintaining equivalent system performance. Experimental results show MEDST outperforms SEDST substantially by 18.7% joint goal accuracy and 14.3% entity match rate on the KVRET corpus with 2% labelled data in semi-supervision.

1 Introduction

Task-oriented dialogue systems (Young et al., 2013) are designed to help users to achieve specific goals such as restaurant reservation or navigation inquiry. In recent years, fully neural end-to-end architectures usually take the sequence-to-sequence (Seq2Seq) (Sutskever et al., 2014) model to generate dialogue responses from the user inputs and context history (Eric and Manning, 2017; Madotto et al., 2018; Wen et al., 2018; Qin et al., 2019; Wu et al., 2019). Neural generative models for task-oriented dialogue systems have achieved promising performance on generation tasks if given a huge training dataset and detailed annotations (Zhao et al., 2017;
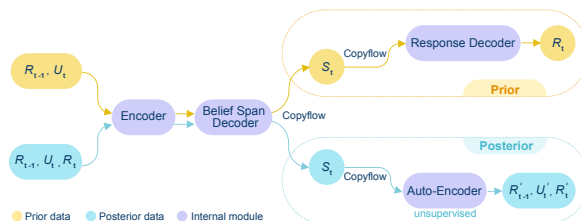

Figure 1: Network architecture of SEDST model.

Lei et al., 2018; Zhang et al., 2019). Arguably, high-quality intermediate labels play a key role in terms of obtaining satisfactory results in this line of tasks (Lei et al., 2018). However, collecting these labels is often the bottleneck in dataset creation, as the process is expensive and time-consuming, requiring domain and expert knowledge (Asri et al., 2017), holding back development in the area of dialogue systems and greatly limiting their application in real-world settings.

Various approaches (Kannan et al., 2018; Tseng et al., 2019; Yin et al., 2019; Chang et al., 2019; Peng et al., 2020) have been proposed recently to conquer this challenge, semi-supervised explicit dialogue state tracker (SEDST) (Jin et al., 2018)[1] is one of them. SEDST tries to solve the label-lacking problem by classical semi-supervision. However, the gain with SEDST is not as satisfactory when the annotated data is very scarce, only a small fraction of what is expected.

In this paper, we focus on improving dialogue performance on top of SEDST. Our first contribution is to enhance SEDST with meta-learning. We propose a MAML-based (Finn et al., 2017) semi-supervision architecture for low-resource. Our experiments with the KVRET dataset show this integration of MAML and SEDST can reach comparable dialogue state tracking accuracy with below 10% intermediate annotation. Second, we improve

---
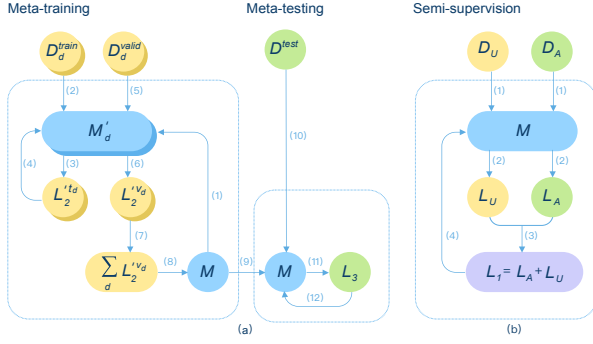
[1] https://github.com/AuCson/SEDST

Figure 2: Yellow indicates the flow of unlabelled data and green indicates labelled data. Three-dimensional graph represents vector constructed of data from D domains ($d = 1, 2, \cdots$), respectively. (a) MEDST contains two phases: meta-training with unlabelled data D_d^{train} and D_d^{valid}, meta-testing with labelled data D^{test}. (b) SEDST model trains the unlabelled data D_U and labelled data D_A in the same phase. More details can be found in our previous work (Huang et al., 2020).

SEDST with Entropy Regularization, that leads to a more robust and better accuracy model. Third, to the best of our knowledge, this proposal is the first attempt to explore meta-based semi-supervised learning for multi-domain task-oriented dialogue tasks, and the novel method can be easily applied to other new scenario too.

2 Proposed Approaches

In this section, we describe the details of MEDST by starting with brief overviews of SEDST.

2.1 SEDST

SEDST is a generative model with copying mechanism and posterior regularization. Dialogue states S_t are represented by text span and flow along the dialogue turns and finally attend to the generation of dialogue response via copying mechanism. Posterior regularization is applied to optimize the training procedure. The normal forward pass of the network calculates the prior probability distribution over vocabulary space which utilize the concatenations of current utterance U_t and previous response R_{t-1} as input. By contrast, the posterior distribution is computed with current actual response R_t added to the inputs which gives more information. The networks will be updated via the KL-divergence of prior and posterior distribution over unlablled data.

Figure 1 illustrates how the process works. It first takes the concatenation of previous responses R_{t-1} and the current user utterance U_t as input and

---

Algorithm 1 MAML based semi-supervision

Input: $D_d^{train}, D_d^{valid}, D^{test}, \delta, \eta, \sigma$.
Output: M.

1: **Meta-training Steps:**
2:   **while** not done **do**
3:     Select unlabelled D_d^t from D_d^{train}, D_d^v from D_d^{valid};
4:     **for** each domain d **do**
5:       Pre-update model with gradient descent:
      $M_d' \leftarrow M - \delta \nabla_M L_2'^{t_d}(M, D_d^t)$
6:       Compute $L_2'^{v_d}(M_d', D_d^v)$ using D_d^v;
7:     **end for**
8:     Update the current model M:
    $M \leftarrow M - \eta \nabla_M \sum_d L_2'^{v_d}(M_d', D_d^v)$
9:   **end while**
10: **Meta-testing Steps:**
11:   **while** not done **do**
12:     Update model with labelled data D^{test}:
    $M \leftarrow M - \sigma \nabla_M L_3(M, D^{test})$
13:   **end while**
14: **return** meta-learned model M;

---

encodes them into hidden vectors. The belief Span decoder is attention-based and extracts the belief span S_t from previous response R_{t-1}, the utterance U_t and previous state S_{t-1}. S_t is then concatenated with R_{t-1} and U_t to generate response R_t. Denote the context as $\mathbf{c} = \{S_{t-1}, R_{t-1}, U_t\}$. The forward propagation network calculates the prior probability distribution P_Θ over vocabulary of S_t. Posterior regularization builds a posterior network which has the same structure with the prior network to learn S_t with the posterior probability distribution Q_Φ.

Posterior regularization is utilized to force prior distribution to approximate that of the posterior network by KL-divergence in training process. While testing, only prior network is active.

In supervised or semi-supervised learning, the learning objectives are defined as,

$$L_1 = - \sum_{}^{A \cup U} \log[P(R_t | R_{t-1}, U_t, S_t)] \quad (1)$$
$$- \sum_{}^{A} \log[P_\Theta(S_t | \mathbf{c}) Q_\Phi(S_t | \mathbf{c}, R_t)]$$
$$+ \alpha \sum_{}^{U} \sum_{i=1}^{N} KL(q_i \| p_i)$$

where A and U indicates the set of labelled and unlabelled data, p_i and q_i are the prior probability and the posterior probability distribution of the i-th

Supervised Proportion	2%		4%		6%		8%	
Metric	Acc	EMR	Acc	EMR	Acc	EMR	Acc	EMR
SEDST	0.546	0.579	0.587	0.605	0.607	0.605	0.699	0.658
MEDST	**0.733**	**0.722**	**0.737**	**0.728**	**0.757**	**0.741**	**0.762**	**0.746**
w/o Entropy	0.688	0.680	0.683	0.667	0.717	0.703	0.710	0.717
w/o MAML	0.556	0.586	0.607	0.609	0.640	0.658	0.709	0.680

Table 1: The semi-supervised evaluation of MEDST and SEDST with different small amounts of labelled data. MEDST achieves better performance in both joint goal accuracy and entity match rate.

word in the state. N is the length of the state span.

When the dataset is entirely unlabelled, the posterior network is extended with an auto-encoder and learning objectives include response generation loss, reconstruction loss and KL-divergence loss,

$$L_2 = -\sum^{U} \log[P(R_t|R_{t-1}, U_t, S_t)] \quad (2)$$
$$-\sum^{U} \log[Q_\Phi(R_t, U_t, R_{t-1}|\hat{S}_t)]$$
$$+\alpha \sum^{U} \sum_{i=1}^{N} KL(q_i\|p_i)$$

2.2 MEDST

We present a new perspective on how to effectively use unlabelled examples for better accuracy and domain adaptation under low-resource. Our proposed model MEDST, motivated by the powerful internal representation ability of meta-learning and the positive effect of entropy in semi-supervised learning, approach the challenges that SEDST remains by the following: (1) MEDST enhances the original loss with entropy regularization. (2) MEDST contains MAML-based semi-supervision on top of SEDST as shown in Algorithm 1.

MEDST includes two phases: meta-training with unlabelled data and meta-testing with labelled data as shown in Figure 2.

In meta-training phase, we train our model similar to SEDST's L_2, further with entropy loss improvement L_2'. As shown in Section 2.1, SEDST suffers data bias when no label resource is available. Following Grandvalet and Bengio (2004), entropy minimization uses a simple loss term to unlabelled data so that the network can make a high confidence (low entropy) prediction. The regularizer can avoid the decision boundary passing through data points which leads to smaller classes overlap.

Therefore, we add the entropy regularization to L_2,

$$L_2' = L_2 - \beta \sum^{U} \sum_{i=1}^{N} p_i \ln(p_i) \quad (3)$$
$$-\beta \sum^{U} \sum_{i=1}^{N} q_i \ln(q_i)$$

In meta-testing phase, a small amount of labelled data is available to optimize the pre-training model. Different from unsupervised learning, labelled data can be utilized to compute the prior and posterior probability distribution of S_t which replaces the entropy of the prior probability p_i and the posterior probability q_i to obtain more deterministic information. The loss function in meta-testing can be derived as,

$$L_3 = L_2 - \lambda \sum^{A} \log[P_\Theta(S_t|\mathbf{c})] \quad (4)$$
$$-\lambda \sum^{A} \log[Q_\Phi(S_t|\mathbf{c}, R_t)]$$

We further explore the adaption ability of MEDST to the new domain with KVRET dataset. Specifically, we choose two domains as source domains with adequate unlabelled data and the other domain as target domain with a small amount of labelled data.

3 Experimental setup

3.1 Corpus and metrics

The KVRET[2] corpus (Eric et al., 2017) is a multi-domain task-oriented dialogue dataset. This dataset includes three distinct domains. There are 284 informable slot values for state tracking. The corpus contains 2425, 302, 302 dialogues for training, validation and testing.

---

[2]http://nlp.stanford.edu/projects/kvret/kvret_dataset_public.zip

Target domain	Supervised Proportion	Generative Model	BLEU	Acc	EMR
weather	5%	SEDST adaption	0.144	0.533	0.446
		MEDST adaption	**0.186**	**0.719**	**0.622**
schedule	5%	SEDST adaption	0.103	0.465	0.457
		MEDST adaption	**0.191**	**0.731**	**0.681**
navigate	5%	SEDST adaption	0.100	0.562	0.689
		MEDST adaption	**0.153**	**0.679**	**0.707**

Table 2: New domain adaption experiment results of MEDST compared to SEDST. MEDST greatly improve the ability of expanding to new domains.
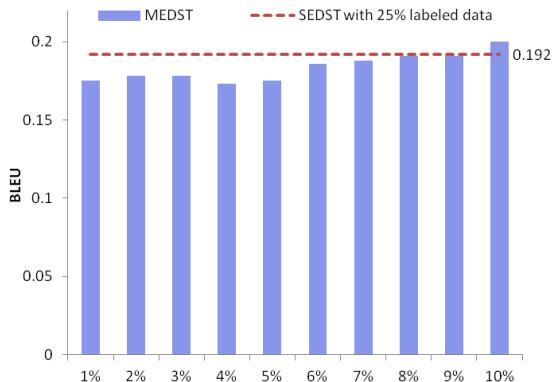

Figure 3: Evaluation of MEDST on BLEU with SEDST using 25% labelled data as baseline. Smaller amount of labelled data can obtain approximate quality of the generated language.

We use three metrics for evaluation following Jin et al. (2018). Joint goal accuracy (Acc) calculates the proportion of the dialogue turns where all the constraints are captured correctly. Entity match rate (EMR) is the proportion that the system captures the correct user goal. We use BLEU (Papineni et al., 2002), a word-overlapping based metric to evaluate the language fluency of generated responses.

3.2 Implementation details

We choose single-layer GRU networks with a hidden size of 50 to be the encoder and the decoder. All the embeddings are initialized by Glove (Pennington et al., 2014) and the size of the word embedding is set to 50. The model is optimized using Adam (Kingma and Ba, 2015) with a learning rate of 0.003 in meta-training and 0.0015 in meta-testing. The model is implemented in the pyTorch.

In MEDST's semi-supervision, we randomly choose 720 unlabelled dialogues for each domain in meta-training with batch size of support dataset 32 and query dataset 8. Parameters β is set to 0.04, α is 0.1 and λ is 1. Different data amounts (2%,

4%, 6%, 8% of training set) of labelled data are utilized for meta-testing and the trained models are selected on the basis of validation performance.

4 Main results

In the main experiments, we take the original SEDST faithful to Jin et al. (2018) as baseline. Table 1 shows MEDST achieves great improvement with different proportions of labelled data. To prove the effectiveness of our structure, we conduct ablation experiments in different setups. w/o Entropy (remove entropy, Acc increases 8.98% and EMR 8.0% on average) has the same regularization loss function as SEDST in meta-training. The improvement here mainly benefits from MAML algorithm, which tries to build an internal representation of multiple tasks and maximize the sensitivity of the loss function when applied to new tasks. w/o MAML (remove MAML, Acc increases 1.83% and EMR 2.15% on average) has the same framework and one-stage training procedure with SEDST. It shows the improvement due to entropy regularization, which takes account of the uncertainty of unlabelled data. We can find that MEDST's advantage mainly comes from MAML and MAML is a potential mechanism in semi-supervision for further studies.

Figure 3 plots another evaluation metric BLEU for MEDST in different amounts of labelled data. MEDST with only 10% labelled data can reach the similar BLEU as SEDST, which requires 25%.

In new domain adaption experiments, our model MEDST performs meta-training on unlabelled data from source domains and meta-testing on 5% labelled data from target domain. SEDST inputs source domain's unlabelled data and 5% target domain's labelled data together to perform one-stage training process. From the results shown in Table 2, we can see MEDST improves the ability of new domain adaption. Three new domains achieve 18.97%

higher in Acc and 13.93% higher in EMR. Target domains can also obtain better generated language quality.

5 Conclusions

In this work, we investigate MAML algorithm and entropy regularization on top of SEDST for low-resource dialogue tasks. We demonstrate the superiority of our proposed model MEDST with low-resource labelled data and perform a fair amount of ablation studies. MEDST can also be adapted to new domains with much better performance. Future work includes exploring more internal mechanism with the combination of semi-supervision and MAML for other tasks.

Acknowledgments

References

Layla El Asri, Hannes Schulz, Shikhar Sharma, Jeremie Zumer, Justin Harris, Emery Fine, Rahul Mehrotra, and Kaheer Suleman. 2017. Frames: a corpus for adding memory to goal-oriented dialogue systems. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 207–219.

Jinxin Chang, Ruifang He, Longbiao Wang, Xiangyu Zhao, Ting Yang, and Ruifang Wang. 2019. A semi-supervised stable variational network for promoting replier-consistency in dialogue generation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pages 1920–1930.

Mihail Eric, Lakshmi Krishnan, Francois Charette, and Christopher D. Manning. 2017. Key-value retrieval networks for task-oriented dialogue. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, Saarbrücken*, pages 37–49.

Mihail Eric and Christopher D. Manning. 2017. A copy-augmented sequence-to-sequence architecture gives good performance on task-oriented dialogue. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pages 468–473.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 1126–1135.

Yves Grandvalet and Yoshua Bengio. 2004. Semi-supervised learning by entropy minimization. In *Advances in Neural Information Processing Systems 17 [Neural Information Processing Systems*, pages 529–536.

Yi Huang, Junlan Feng, Min Hu, Xiaoting Wu, Xiaoyu Du, and Shuo Ma. 2020. Meta-reinforced multi-domain state generator for dialogue systems. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7109–7118.

Xisen Jin, Wenqiang Lei, Zhaochun Ren, Hongshen Chen, Shangsong Liang, Yihong Zhao, and Dawei Yin. 2018. Explicit state tracking with semi-supervision for neural dialogue generation. *Computing Research Repository*, arXiv:1808.10596.

Anjuli Kannan, Kai Chen, Diana Jaunzeikare, and Alvin Rajkomar. 2018. Semi-supervised learning for information extraction from dialogue. In *Interspeech 2018, 19th Annual Conference of the International Speech Communication Association*, pages 2077–2081.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations*.

Wenqiang Lei, Xisen Jin, Min-Yen Kan, Zhaochun Ren, Xiangnan He, and Dawei Yin. 2018. Sequicity: Simplifying task-oriented dialogue systems with single sequence-to-sequence architectures. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 1437–1447.

Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018. Mem2seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 1468–1478.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318.

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. *CoRR*, abs/2002.12328.

Jeffrey Pennington, Richard Socher, and Christopher D. Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 1532–1543.

Libo Qin, Yijia Liu, Wanxiang Che, Haoyang Wen, Yangming Li, and Ting Liu. 2019. Entity-consistent end-to-end task-oriented dialogue system with KB retriever. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*

and the 9th International Joint Conference on Natural Language Processing, pages 133–142.

Ilya Sutskever, Oriol Vinyals, and Quoc V. Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems*, pages 3104–3112.

Bo-Hsiang Tseng, Marek Rei, Pawel Budzianowski, Richard E. Turner, Bill Byrne, and Anna Korhonen. 2019. Semi-supervised bootstrapping of dialogue state trackers for task-oriented modelling. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, pages 1273–1278.

Haoyang Wen, Yijia Liu, Wanxiang Che, Libo Qin, and Ting Liu. 2018. Sequence-to-sequence learning for task-oriented dialogue with dialogue state representation. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3781–3792.

Chien-Sheng Wu, Richard Socher, and Caiming Xiong. 2019. Global-to-local memory pointer networks for task-oriented dialogue. In *7th International Conference on Learning Representations*.

Yichun Yin, Lifeng Shang, Xin Jiang, Xiao Chen, and Qun Liu. 2019. Dialog state tracking with reinforced data augmentation. *CoRR*, abs/1908.07795.

Steve J. Young, Milica Gasic, Blaise Thomson, and Jason D. Williams. 2013. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.

Yichi Zhang, Zhijian Ou, and Zhou Yu. 2019. Task-oriented dialog systems that consider multiple appropriate responses under the same context. *CoRR*, abs/1911.10484.

Tiancheng Zhao, Allen Lu, Kyusong Lee, and Maxine Eskénazi. 2017. Generative encoder-decoder models for task-oriented spoken dialog systems with chatting capability. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 27–36.