

Cost-Sensitive Active Learning for Dialogue State Tracking

Kaige Xie, Cheng Chang, Liliang Ren, Lu Chen and Kai Yu

Key Lab. of Shanghai Education Commission for Intelligent Interaction and Cognitive Eng.

SpeechLab, Department of Computer Science and Engineering

Brain Science and Technology Research Center

Shanghai Jiao Tong University, Shanghai, China

{lightyear0117, cheng.chang, ren11204, chenlusz, kai.yu}@sjtu.edu.cn

Abstract

Dialogue state tracking (DST), when formulated as a supervised learning problem, relies on labelled data. Since dialogue state annotation usually requires labelling all turns of a single dialogue and utilizing context information, it is very expensive to annotate all available unlabelled data. In this paper, a novel *cost-sensitive* active learning framework is proposed based on a set of new dialogue-level query strategies. This is the first attempt to apply active learning for dialogue state tracking. Experiments on DSTC2 show that active learning with mixed data query strategies can effectively achieve the same DST performance with significantly less data annotation compared to traditional training approaches.

1 Introduction

The dialogue state tracker, an important component of a spoken dialogue system, tracks the internal belief state of the system based on the history of the dialogue. For each turn, the tracker outputs a distribution over possible dialogue states, on which the dialogue system relies to take proper actions to interact with users. Various approaches have been proposed for dialogue state tracking, including *hand-crafted rules* (Wang and Lemon, 2013; Sun et al., 2014), *generative models* (Thomson and Young, 2010; Young et al., 2010, 2013) and *discriminative models* (Ren et al., 2013; Lee and Eskenazi, 2013; Williams, 2014). For *discriminative models*, recent studies on data-driven approaches have shown promising performance, especially on Recurrent Neural Network (RNN) (Henderson et al., 2013, 2014c). As for datasets, the Dialogue State Tracking Chal-

lenge (DSTC) series (Williams et al., 2016) have provided common testbeds for this task.

Though data-driven approaches have achieved promising performance, they require large amounts of labelled data, which are costly to be fully annotated. Besides this, it is quite difficult to label a single dialogue because, for every dialogue turn, experts need to label all the semantic slots and typically, to label a single turn accurately, they need to pay attention to the context rather than the current turn only. Active learning (AL) (Settles, 2010) can be applied to select valuable samples to label. Using the AL approach, we need fewer labelled samples when training the model to reach the same or even better performance compared to traditional training approaches.

Although it is often assumed that the labelling costs are the same for all samples in some tasks (González-Rubio and Casacuberta, 2014; Sivaraman and Trivedi, 2014), it is appropriate to consider different labelling costs for the dialogue state tracking task where different dialogues vary in the number of turns. In this paper, we define the labelling cost for each dialogue sample with respect to its number of dialogue turns. Then we provide a new AL query criterion called *diversity*, and finally propose a novel *cost-sensitive* active learning approach based on three dimensions: *cost*, *uncertainty*, and *diversity*. The results of experiments on the DSTC2 dataset (Henderson et al., 2014a) demonstrate that our approaches are more effective compared to traditional training methods.

In the next section, we will present the proposed cost-sensitive active learning framework for dialogue state tracking. Then in Section 3 we will describe the experimental setup and show the results of experiments on the DSTC2 dataset, followed by our conclusions and future work in Section 4.

2 Cost-Sensitive Active Learning

A complete work cycle of active learning for dialogue state tracking includes 3 steps: (1) train the tracker with labelled dialogue samples; (2) post query using the query strategy to select the valuable unlabelled dialogue and ask experts for its label; (3) merge the newly-labelled dialogue with all previously-labelled dialogue samples and return to (1). The tracker and query strategy will be introduced in Section 2.1 and 2.2 respectively.

2.1 Dialogue State Tracker

Our proposed active learning workflow is independent of the tracker type. Here we use the *LecTrack* model (Zilka and Jurcicek, 2015) as a word-based tracker. For each turn t in a dialogue, the tracker takes in a word concatenation of all history words (together with their confidence scores from ASR) within this dialogue and finally outputs a prediction. The general model structure (at turn t) is shown in Figure 1. The notation $\mathbf{w} \oplus \mathbf{u}$ denotes the concatenation of two vectors, \mathbf{w} (the word) and \mathbf{u} (the confidence score). FC refers to the Fully Connected Layer. The output of FC is then encoded by the LSTM encoder Enc, whose output (only the last one) will be sent to a Softmax layer to make a prediction $\mathbf{p}_t^s \in \mathbb{R}^{N_s}$ over all N_s possible values for a given slot s at turn t :

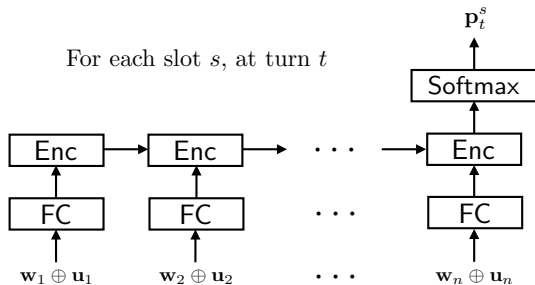


Figure 1: *LecTrack* model structure. 4 models in total since $s \in \{food, pricerange, name, area\}$.

2.2 Cost-Sensitive Active Learning Methods

Given that different dialogues vary in the number of turns, we assume that the smallest query unit should be a whole dialogue, and that the cost for labelling a dialogue is directly proportional to the number of dialogue turns.

Since all the unlabelled data is possible to be collected simultaneously, the DST task can be regarded as *pool-based sampling* (Settles, 2010). This assumes that there is a small pool of labelled data \mathcal{L} , and a large pool of unlabelled data \mathcal{U} available. That allows us to query the samples in a

greedy fashion according to some measurement criteria, which are used to evaluate all samples in the unlabelled pool.

We propose four novel query strategies. The first three utilize one kind of measurement criterion respectively and the last one is based on the mixture of the first three. For convenience, a certain dialogue sample is denoted as x .

Cost Strategy (CS)

This strategy prefers the dialogue samples that have the minimum number of turns. For each dialogue sample, its labelling cost, denoted as $C(x)$, can be defined as the number of turns.

Uncertainty Strategy (US)

This strategy prefers the dialogue samples whose predictions the DST model is most uncertain about. In this paper, we take advantage of *entropy* (Shannon, 2001) as the uncertainty measurement criterion. The dialogue uncertainty on slot s , $U_s(x)$, is the average over all the *entropy* values of dialogue turns:

$$U_s(x) = -\frac{1}{C(x)} \sum_{t=1}^{C(x)} \sum_{i=1}^{N_s} \mathbf{p}_t^s[i] \log \mathbf{p}_t^s[i],$$

where \mathbf{p}_t^s can be directly obtained from the DST model described in Section 2.1.

Diversity Strategy (DS)

This strategy prefers the dialogue samples that are most diverse from the dialogues currently in the labelled pool \mathcal{L} . As the training and querying process goes on, the diversity of dialogue samples selected to be labelled will decrease gradually, which results in a biased training process. To handle such problem, here we design a novel Spherical k -Means Clustering (MacQueen et al., 1967) based method to evaluate the diversity of dialogue samples and select the most diverse ones in unlabelled pool \mathcal{U} to label, so that we could maintain the diversity of dialogue samples in labelled pool \mathcal{L} .

Different dialogues have varying lengths so an embedding function to map each dialogue into a fixed-dimensional continuous space is needed. We utilize the method of *unsupervised dialogue embeddings* (Su et al., 2016) to extract a dialogue feature, which is used to calculate the diversity.

We choose the bag-of-words (BOW) representation as a turn-level feature \mathbf{f}_t at turn t , which will be sent into a Bi-directional

LSTM (BLSTM) (Graves et al., 2013) encoder to obtain the two directional hidden sequences, $\mathbf{h}_{1:T}^f$ and $\mathbf{h}_{1:T}^b$. At turn t , $\mathbf{h}_t^f = LSTM^f(\mathbf{f}_t, \mathbf{h}_{t-1}^f)$ and $\mathbf{h}_t^b = LSTM^b(\mathbf{f}_t, \mathbf{h}_{t+1}^b)$. Then, the dialogue feature vector \mathbf{v} is calculated as the average over all hidden sequences, i.e., $\mathbf{v} = \frac{1}{T} \sum_{t=1}^T \mathbf{h}_t^f \oplus \mathbf{h}_t^b$, where the notation $\mathbf{h}_t^f \oplus \mathbf{h}_t^b$ denotes the concatenation of the two vectors, \mathbf{h}_t^f and \mathbf{h}_t^b .

Next, the dialogue feature vector is chosen as the input of a forward LSTM decoder for each turn t , which ultimately outputs feature sequences $\mathbf{f}'_{1:T}$. The model’s training target is to minimize the mean-square-error (MSE) between $\mathbf{f}_{1:T}$ and $\mathbf{f}'_{1:T}$.

The feature vectors of all dialogues in both \mathcal{L} and \mathcal{U} can be obtained with this pre-trained model. Define $\mathcal{V}_{\mathcal{L}} = \{\mathbf{v}_{l_1}, \mathbf{v}_{l_2}, \dots\}$ as the feature vector set of \mathcal{L} and $\mathcal{V}_{\mathcal{U}} = \{\mathbf{v}_{u_1}, \mathbf{v}_{u_2}, \dots\}$ as the feature vector set of \mathcal{U} .

We fit the set $\mathcal{V}_{\mathcal{L}}$ into a Spherical k -Means model with N_c clusters, so that we can acquire a substitutional set of feature vectors denoted as $\mathcal{V}'_{\mathcal{L}} = \{\mathbf{v}'_{l_1}, \mathbf{v}'_{l_2}, \dots, \mathbf{v}'_{l_{N_c}}\}$, which is composed of N_c representative feature vectors (clusters) among the vectors in the original set $\mathcal{V}_{\mathcal{L}}$. Then for each vector \mathbf{v}_{u_i} in $\mathcal{V}_{\mathcal{U}}$, calculate its *cosine similarity* to N_c vectors in $\mathcal{V}'_{\mathcal{L}}$ respectively, and regard the maximum of N_c similarity values as its true similarity to the whole labelled set, since the cluster of maximum similarity has the largest representativeness of all the original vectors in the labelled set. Therefore, the diversity measure $D(x)$ can be defined as the opposite number of similarity:

$$D(x) = - \max_{i=1, \dots, N_c} \left\{ \frac{\mathbf{v}_{u_x} \cdot \mathbf{v}'_{l_i}}{\|\mathbf{v}_{u_x}\| \cdot \|\mathbf{v}'_{l_i}\|} \right\}.$$

Mixed Strategy

In practice, we usually need different query strategies at different learning stages (Settles, 2010). Based on the strategies presented above, we propose a new query strategy called **Cost-Uncertainty-Diversity Strategy (CUDS)**, which is originated from the idea of combining multiple strategies. This strategy takes into consideration three measurement criteria, i.e. *cost*, *uncertainty* and *diversity*, so that the unlabelled samples can be evaluated more comprehensively.

Specifically, what we want is to select samples with low cost, high uncertainty and high diversity. Based on this, we propose a new measurement criterion, denoted as $M(x)$. Naturally, the goal of **CUDS** is to pick out the dialogue samples which

have the maximum measurement value $M(x)$:

$$M(x) = -\alpha C(x) + \beta U_s(x) + \gamma D(x), \quad (1)$$

where α , β and γ are positive weighting parameters that can be tuned so as to find a good trade-off among those three measurement criteria.

In order to conduct weighting, $C(x)$, $U_s(x)$ and $D(x)$ should possess the same scale. $C(x)$ ranges from 1 to C_{max} , the maximum number of a single dialogue’s turns, and therefore we replace $C(x)$ in Equation 1 with $C_m(x) = C(x)/C_{max}$. The range of $D(x)$ is $(-1, 1)$, so we replace $D(x)$ in Equation 1 with $D_m(x) = (D(x)+1)/2$. Then the original Equation 1 is transformed into Equation 2:

$$M(x) = -\alpha C_m(x) + \beta U_s(x) + \gamma D_m(x). \quad (2)$$

3 Experiments and Results

3.1 Experimental Setup

Experiments are conducted to assess the performance of different query strategies on single slot and joint goal respectively. The dataset we use is the second Dialogue State Tracking Challenge (DSTC2) dataset (Henderson et al., 2014a), which focuses on the restaurant information domain and contains 7 slots of which 4 are informable and all 7 are requestable. We implement the dialogue state tracker as described in Section 2.1. The model is trained using Stochastic Gradient Descent (SGD), collaborating with a gradient clipping heuristic (Pascanu et al., 2012) to avoid the exploding gradient problems.

3.2 Results on Single Slot

In this section, five different query strategies are compared on single slot. Besides the four query strategies presented in Section 2.2, here we choose **Random Strategy (RS)** as our baseline query strategy. **RS** means we randomly select dialogues to annotate. Although it may seem quite simple, we have to point out that such naive strategy does perform not bad in practice. We attribute such phenomenon to the fact that the query process is dominated by the underlying distribution of the original dataset. A nature of AL called *sampling bias* (Dasgupta and Hsu, 2008) can be considered as the main cause. The training set may gradually diverge from the real data distribution as the training and querying process continues. However, **RS** is luckily not influenced by this effect, which allows it to be a powerful baseline to compare with.

According to the current strategy, the model

queries 2 dialogue samples each time. Figure 2 displays the training accuracy curves of the *food* slot (the most difficult slot) using different query strategies. Here we use the number of labelled dialogue turns as the x -axis, which can be regarded as the labelling cost. It is shown that the three query strategies (**CS/US/DS**), which are based on single measurement criterion respectively, have better performance than the baseline strategy (**RS**). The reason why **DS** does not perform very well at the beginning is that the diverse but greatly scarce data is not sufficient to train an effective model. Our proposed mixed strategy **CUDS** achieves the best performance among all the query strategies, which proves the effectiveness of our strategy mixture methodology. Considering the training cost, although the DSTC2 training set is composed of **11677** turns (**1612** dialogues) in total, **CUDS** only consumes about **3000** turns (about **520** dialogues) for training to convergence. Besides, while **RS** consumes **5000** turns (about **700** dialogues) when converging, **CUDS** just consumes **2000** turns (about **360** dialogues) to achieve the performance equal to the convergence level of **RS**, reducing the cost by **60%**.

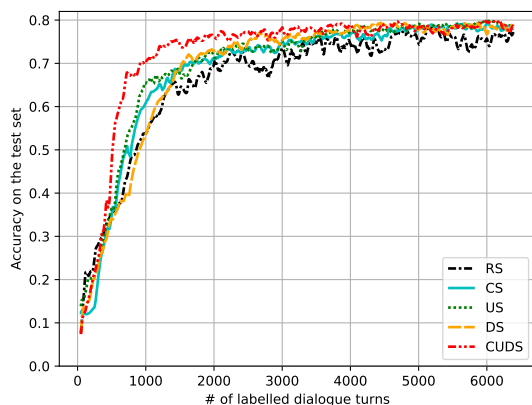


Figure 2: Curves of *food* slot.

3.3 Results on Joint Goal

Figure 3 displays the training accuracy curves of the joint goal using five different query strategies. At different learning stages, the query strategy of best performance is different. **US** rises more quickly at the beginning while **DS** diverges earlier. The reasons include: in order to finally reach convergence, the tracker need to see samples of great diversity, which allows it to give consideration to several semantic slots; however, samples of large *entropy* can bring tracker more concrete information on the most controversial cases, which helps it to learn from scratch rapidly. Our mixed strategy

CUDS, combining the advantages of those two while minimizing the cost at the same time, obtains a performance improvement. Although the final convergence level is not quite high due to the limitation of *LecTrack* model, it does not diminish the effectiveness of proposed AL query strategies.

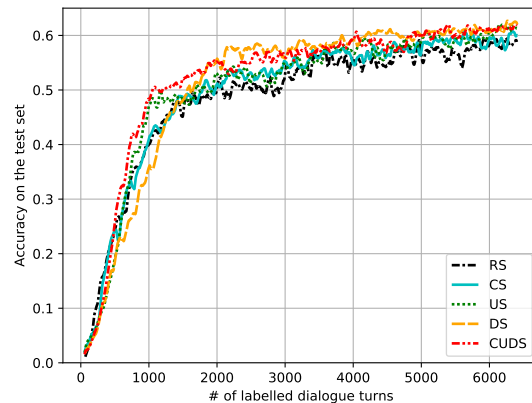


Figure 3: Curves of joint goal.

4 Conclusions and Future Work

In this paper, a novel cost-sensitive active learning technique is presented for dialogue state tracking, which assumes that each dialogue sample has a nonuniform labelling cost related to the number of dialogue turns. Besides *cost*, we also provide another two measurement criteria, *uncertainty* and *diversity*. Our mixed query strategy considers those three criteria comprehensively in order to make queries more appropriately. Experiment results demonstrate that our proposed approaches can achieve promising tracking performance with lower cost compared to traditional methods.

Our future work roughly includes two parts. One is to deploy our proposed AL methods on some other dialogue tasks such as DSTC3 (Henderson et al., 2014b) to verify the results presented in this paper. The other is to conduct our approaches on DST models of better performance (Mrkšić et al., 2017) because the model’s tracking ability has an inevitable influence on the whole active learning process.

Acknowledgements

The corresponding author is Kai Yu. This work has been supported by Shanghai International Science and Technology Cooperation Fund (No. 16550720300), the Major Program of Science and Technology Commission of Shanghai Municipality (STCSM) (No. 17JC1404104) and the JiangSu NSFC project (BE2016078).

References

- Sanjoy Dasgupta and Daniel Hsu. 2008. Hierarchical sampling for active learning. In *Proceedings of the 25th international conference on Machine learning*, pages 208–215. ACM.
- Jesús González-Rubio and Francisco Casacuberta. 2014. Cost-sensitive active learning for computer-assisted translation. *Pattern Recognition Letters*, 37:124–134.
- Alex Graves, Navdeep Jaitly, and Abdel-rahman Mohamed. 2013. Hybrid speech recognition with deep bidirectional lstm. In *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*, pages 273–278. IEEE.
- Matthew Henderson, Blaise Thomson, and Jason D Williams. 2014a. The second dialog state tracking challenge. In *Proceedings of the SIGDIAL 2014 Conference*, pages 263–272.
- Matthew Henderson, Blaise Thomson, and Jason D Williams. 2014b. The third dialog state tracking challenge. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*, pages 324–329. IEEE.
- Matthew Henderson, Blaise Thomson, and Steve Young. 2013. Deep neural network approach for the dialog state tracking challenge. In *Proceedings of the SIGDIAL 2013 Conference*, pages 467–471.
- Matthew Henderson, Blaise Thomson, and Steve Young. 2014c. Robust dialog state tracking using delexicalised recurrent neural networks and unsupervised adaptation. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*, pages 360–365. IEEE.
- Sungjin Lee and Maxine Eskenazi. 2013. Recipe for building robust spoken dialog state trackers: Dialog state tracking challenge system description. In *Proceedings of the SIGDIAL 2013 Conference*, pages 414–422.
- James MacQueen et al. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.
- Nikola Mrkšić, Diarmuid Ó Séaghdha, Tsung-Hsien Wen, Blaise Thomson, and Steve Young. 2017. Neural belief tracker: Data-driven dialogue state tracking. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 1777–1788.
- Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. 2012. Understanding the exploding gradient problem. *CoRR*, abs/1211.5063.
- Hang Ren, Weiqun Xu, Yan Zhang, and Yonghong Yan. 2013. Dialog state tracking using conditional random fields. In *Proceedings of the SIGDIAL 2013 Conference*, pages 457–461.
- Burr Settles. 2010. Active learning literature survey. *University of Wisconsin, Madison*, 52(55-66):11.
- Claude E Shannon. 2001. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55.
- Sayanan Sivaraman and Mohan M Trivedi. 2014. Active learning for on-road vehicle detection: A comparative study. *Machine vision and applications*, pages 1–13.
- Pei-Hao Su, Milica Gasic, Nikola Mrkšić, Lina M Rojas Barahona, Stefan Ultes, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2016. On-line active reward learning for policy optimisation in spoken dialogue systems. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 2431–2441.
- Kai Sun, Lu Chen, Su Zhu, and Kai Yu. 2014. A generalized rule based tracker for dialogue state tracking. In *Spoken Language Technology Workshop (SLT), 2014 IEEE*, pages 330–335. IEEE.
- Blaise Thomson and Steve Young. 2010. Bayesian update of dialogue state: A pomdp framework for spoken dialogue systems. *Computer Speech & Language*, 24(4):562–588.
- Zhuoran Wang and Oliver Lemon. 2013. A simple and generic belief tracking mechanism for the dialog state tracking challenge: On the believability of observed information. In *Proceedings of the SIGDIAL 2013 Conference*, pages 423–432.
- Jason Williams, Antoine Raux, and Matthew Henderson. 2016. The dialog state tracking challenge series: A review. *Dialogue & Discourse*, 7(3):4–33.
- Jason D Williams. 2014. Web-style ranking and slu combination for dialog state tracking. In *Proceedings of the SIGDIAL 2014 Conference*, pages 282–291.
- Steve Young, Milica Gašić, Simon Keizer, François Mairesse, Jost Schatzmann, Blaise Thomson, and Kai Yu. 2010. The hidden information state model: A practical framework for pomdp-based spoken dialogue management. *Computer Speech & Language*, 24(2):150–174.
- Steve Young, Milica Gašić, Blaise Thomson, and Jason D Williams. 2013. Pomdp-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.
- Lukas Zilka and Filip Jurcicek. 2015. Incremental lstm-based dialog state tracker. In *Automatic Speech Recognition and Understanding (ASRU), 2015 IEEE Workshop on*, pages 757–762. IEEE.