PROGRESSIVE DIALOGUE STATE TRACKING FOR MULTI-DOMAIN DIALOGUE SYSTEMS

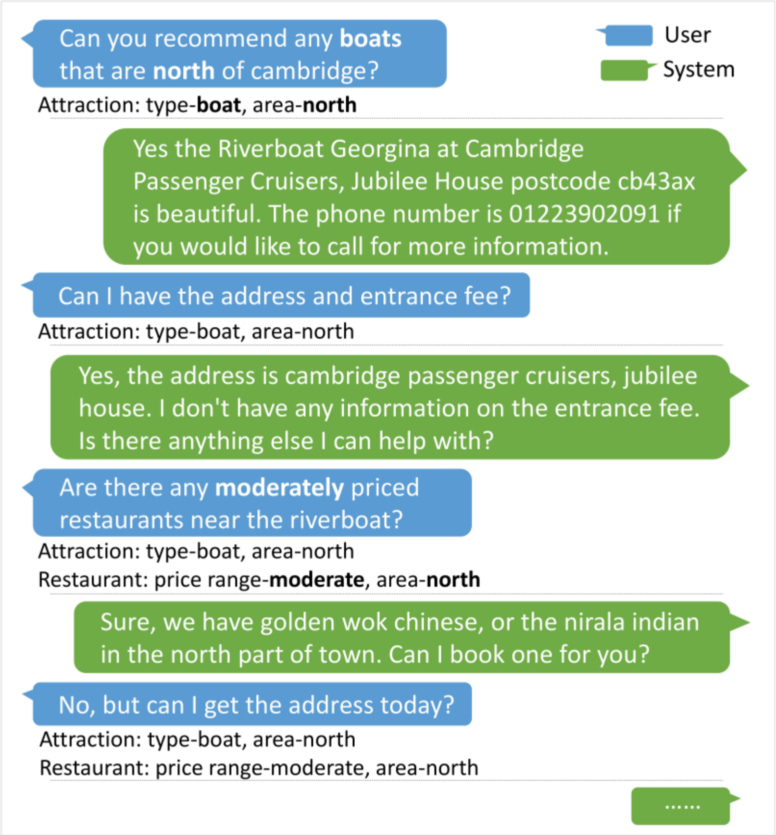
ICASSP 2021

Jiahao Wang[*], Minqian Liu[†] and Xiaojun Quan[*]

[*]Sun Yat-sen University
[†]South China University of Technology

1. Background

Dialogue State Tracking (Multi-Domain)



Dialogue states → Domain-Slot-Value

Turn 0:
- **attraction-type-boat,**
- **attraction-area-north**

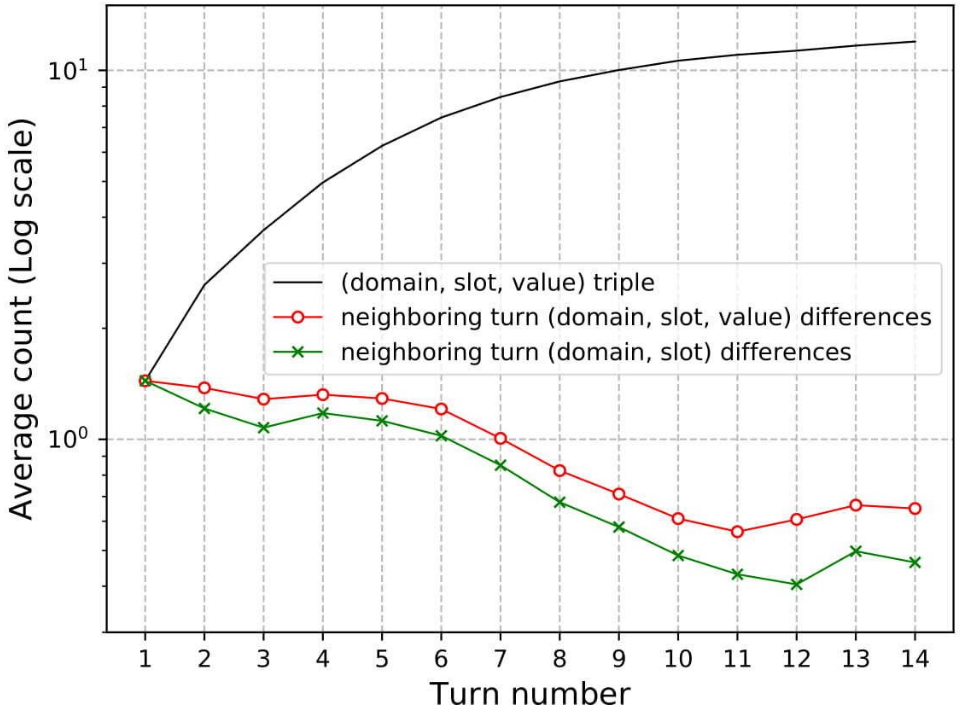
Turn 1:
- attraction-type-boat,
- attraction-type-north

Turn 2:
- attraction-type-boat,
- attraction-type-north,
- **restaurant-price range-moderate,**
- **restaurant-area-north**

Turn 3:
- attraction-type-boat,
- attraction-type-north,
- restaurant-price range-moderate,
- restaurant-area-north

2. Motivation

Two observations in MultiWOZ dataset



- The domain-slot-value triples increase with turns
- Differences of domain-slot-value triples between neighboring turns are small

2. Motivation

Two observations in MultiWOZ dataset -- Formalization

Define:

- $B^{(t)} = \{b_1^{(t)}, b_2^{(t)}, \ldots, b_m^{(t)}\}$: Dialogue state at the t-th turn
- $b_i^{(t)}$: i-th domain-slot pair with label from $\{pointed, dontcare, none\}$
- $\nabla^{(t)} = \{\delta_1^{(t)}, \delta_2^{(t)}, \ldots, \delta_m^{(t)}\}$: Changes from $B^{(t-1)}$ to $B^{(t)}$

$$\delta_i^{(t)} = f(x) = \begin{cases} b_i^{(t)}, & b_i^{(t)} \neq b_i^{(t-1)} \\ \emptyset, & b_i^{(t)} = b_i^{(t-1)} \end{cases}$$

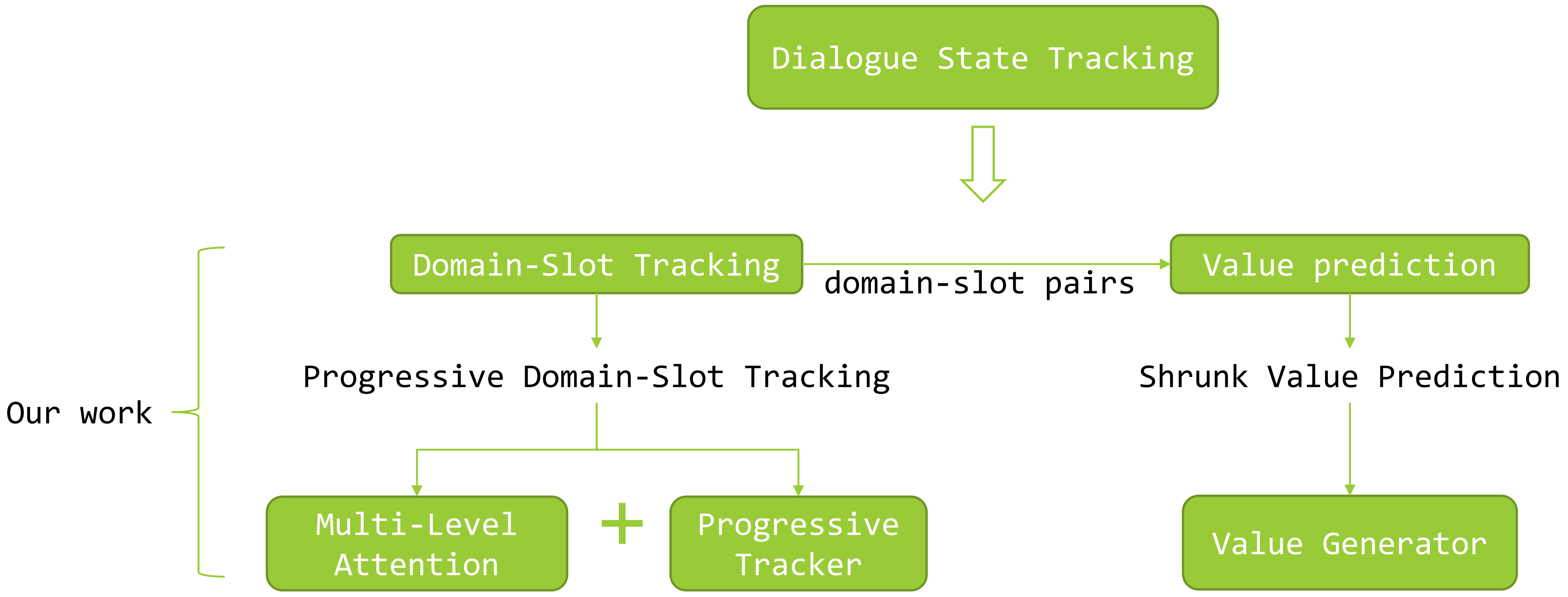
- Accumulating state triples:

  For most utterances: $|B_{pointed}^{(t-1)}| < |B_{pointed}^{t}|$

- Adjacent state dependencies:

  $\Delta^{(t)}$ is small

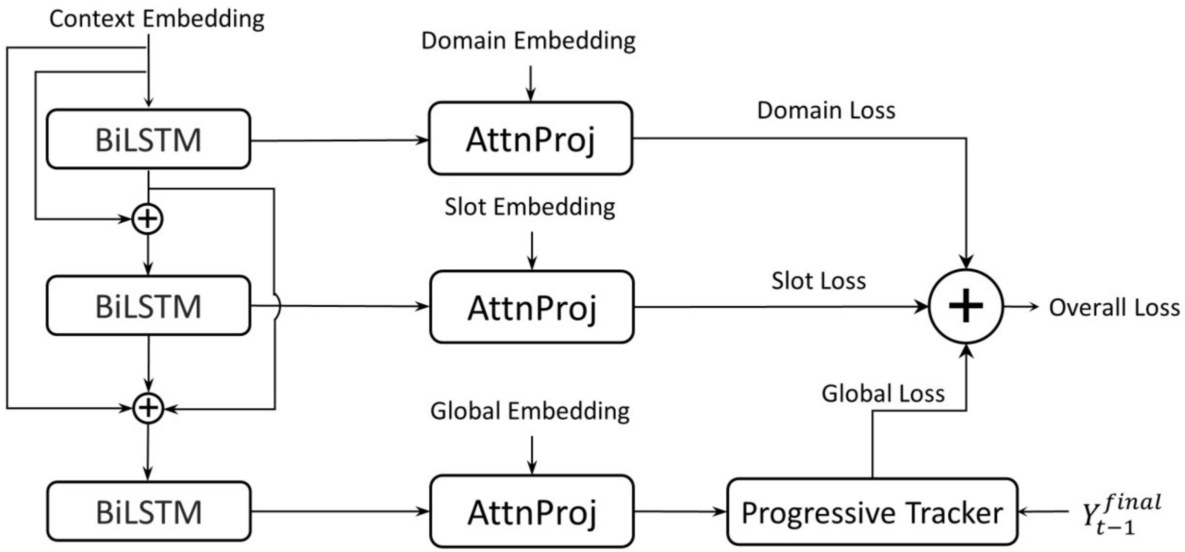
3. Method


Dialogue State Tracking
Domain-Slot Tracking
domain-slot pairs
Value prediction
Progressive Domain-Slot Tracking
Shrunk Value Prediction
Our work
Multi-Level Attention
+
Progressive Tracker
Value Generator

3. Method

Progressive Domain-Slot Tracking

Multi-Level Attention

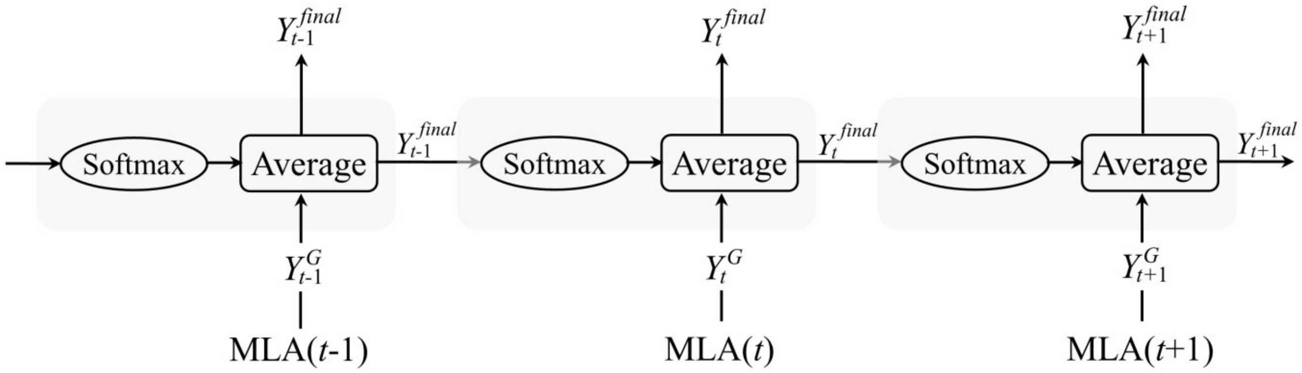


- We adopt the attention mechanism on domain, slot, and global (combination of domains and slots) levels
  - We use three levels of attention to capture information of different levels
  - We compute the loss of each level separately to help optimize the module

3. Method

Progressive Domain-Slot Tracking

Progressive Tracker



$$Y_t^{final} = \frac{\mathrm{Softmax}(Y_{t-1}^{final}) + Y_t^G}{2}$$

Inspired by the two observations:

- We design a progressive domain-slot tracker to make use of the predictions from the previous turn
  - Progressive tracking allows the model to focus more on the change between adjacent turns
  - We adopt a Softmax layer as a smoothing operation to reduce the risk of error propagation

4. Results

Joint Goal Accuracy

Model	**MultiWOZ 2.0 Acc. (%)**	**MultiWOZ 2.1 Acc. (%)**
GLAD [2]	35.57	-
DST Reader [17]	39.41	36.4
COMER [9]	45.72	-
TRADE [8]	48.62	45.6
NADST [18]	50.52	49.0
SAS [19]	51.03	-
DST-SC [20]	**52.24**	49.6
PRO-DST (Ours)	51.48	**49.9**

- Experiments on MultiWOZ show that our method achieves promising results

4. Results

Ablation Study

Row	Model	Domain-Slot Acc.	Goal Acc.
1	Our Model	58.06	49.89
2	− output of previous turn	56.28	48.4
3	− smoothing of previous-turn output	51.14	44.75
4	− domain&slot attention module	47.92	41.99
5	− learnable domain&slot emb. + fixed domain&slot emb.	58.06	49.80

- The predictions in the previous turn effectively improve the performance, but the smoothing operation is critical
- The information in the domain and slot levels improves the performance significantly

5. Discussion

- The two observations can direct further research on developing more accurate domain-slot tracker, e.g., utilizing large-scaled pretrained language model.
- The three-level attentions enables finer-grained modeling of domain-slot predictions and can be extended to more complicate dialogue state ontology setting.
- The progressive domain-slot tracking mechanism can be improved to focus more on the state changes between neighboring turns.

Thank you