

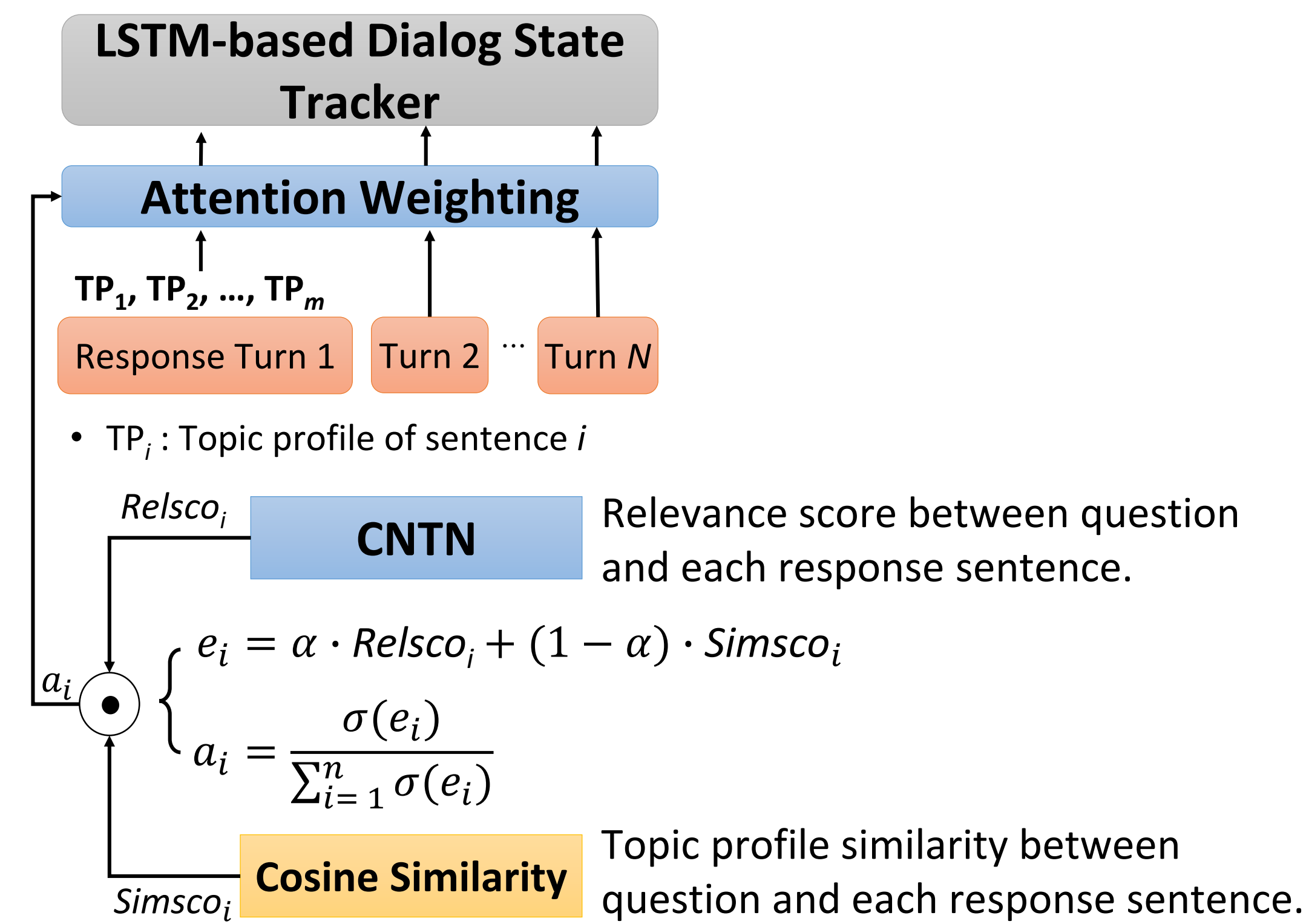
Attention-based Dialog State Tracking for Conversational Interview Coaching

Ming-Hsiang Su, Chung-Hsien Wu, Kun-Yi Huang and Chu-Kwang Chen

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan, TAIWAN

Introduction

- **Goal:** Propose an **attention weighting mechanism** for **dialog state tracking** in a conversational interview coaching system.
- **Problem:**
 1. The **semantic slots** for dialog state representation are difficult to define manually.
 2. The interviewee's response contains **irrelevant sentences**.
- **Approach:**
 1. **Topic profile similarity** between question and a response sentence is estimated based on **LDA-based topic model**.
 2. **Relevance score** between question and a response sentence is obtained based on **convolutional neural tensor network (CNTN)**.
 3. **Attention mechanism** for **sentence weighting**.
 4. **Two-layer LSTM-based autoencoder** for **dialog state tracking**.

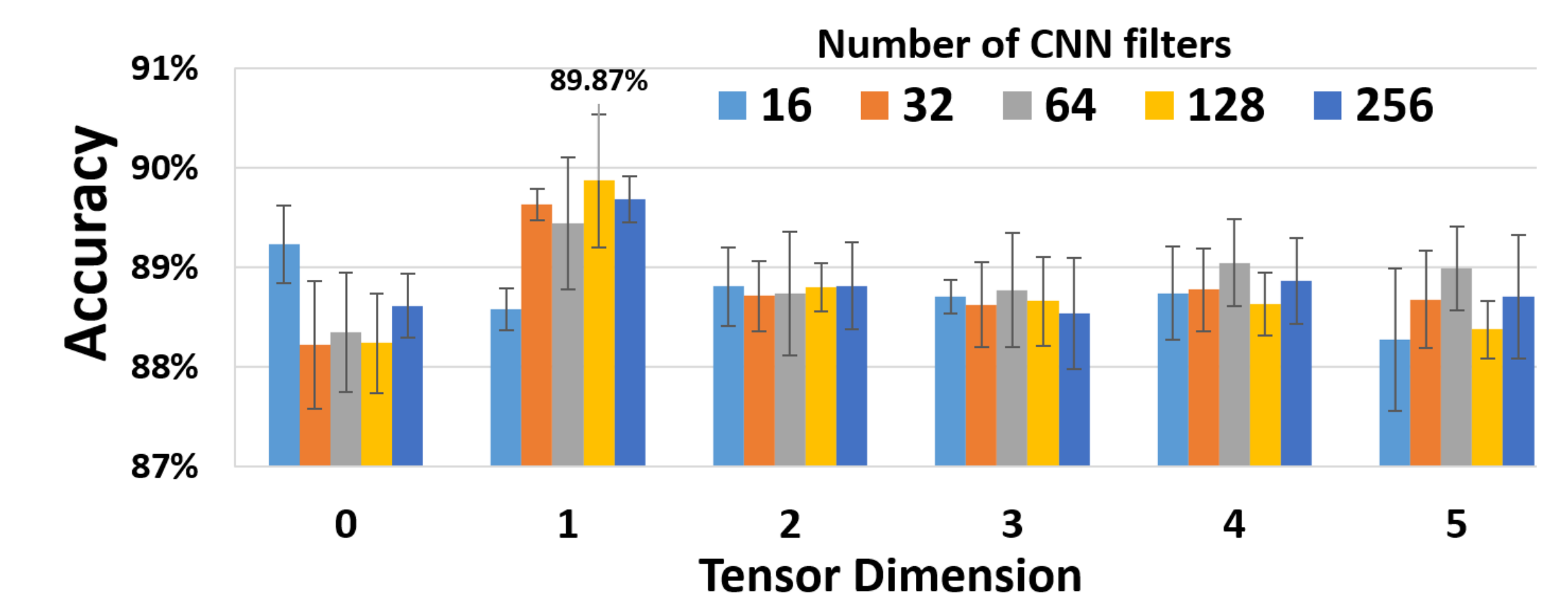


Experimental Results

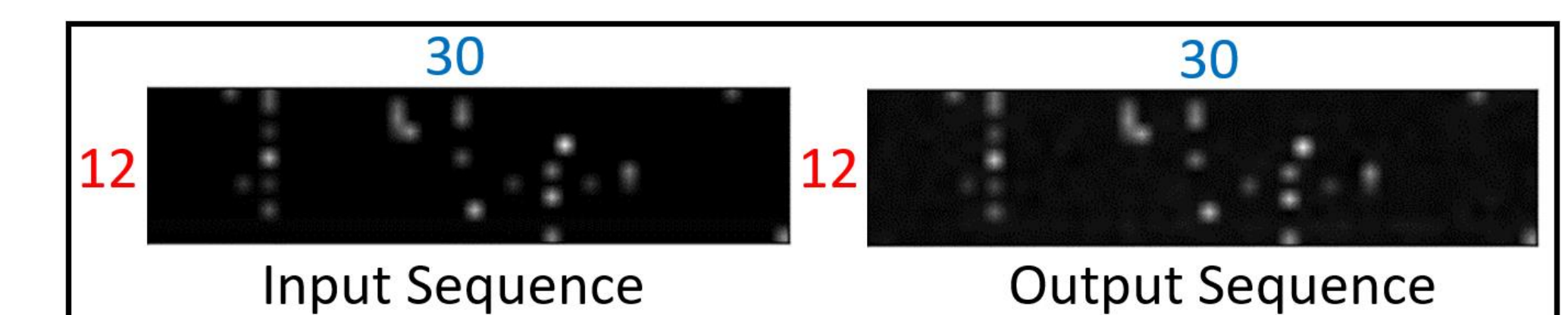
- **NCKU Interview Database** collected **260** dialogs with **1754** ordinary questions and **1262** follow-up questions.

	Total
Number of turns	3016
Average number of turns	10.7
Average number of sentences in each answer	3.84
Interview time (minutes) per interview	20

- The **accuracy of CNTN-based attention model** achieved **89.87%**, when the tensor dimensionality was 1 and the number of CNN filters was 128



- **LSTM-based autoencoder** was evaluated by visualizing the vector representation of input and output sequence, when the number of **topics** was **30**.



- **Performance Comparison** between the traditional and the proposed methods using the same Double Q-learning agent.

Method	Semantic slot (baseline)	Topic profile	
		w/o sentence attention	w/ sentence attention
$\alpha = 0.5$			
Parameters	10 semantic slots Dialog states = 256	30-D topic profile Dialog states = 64	10-D topic profile Dialog states = 128
Turn	4.74	4.80	5.72
Diff	1.00	0.94	0.02
Accumulative Reward	6.46	6.34	7.27

Conclusions

- A five-fold cross validation scheme was employed and the results show that the proposed method outperformed the semantic slot-based method.

System Framework

