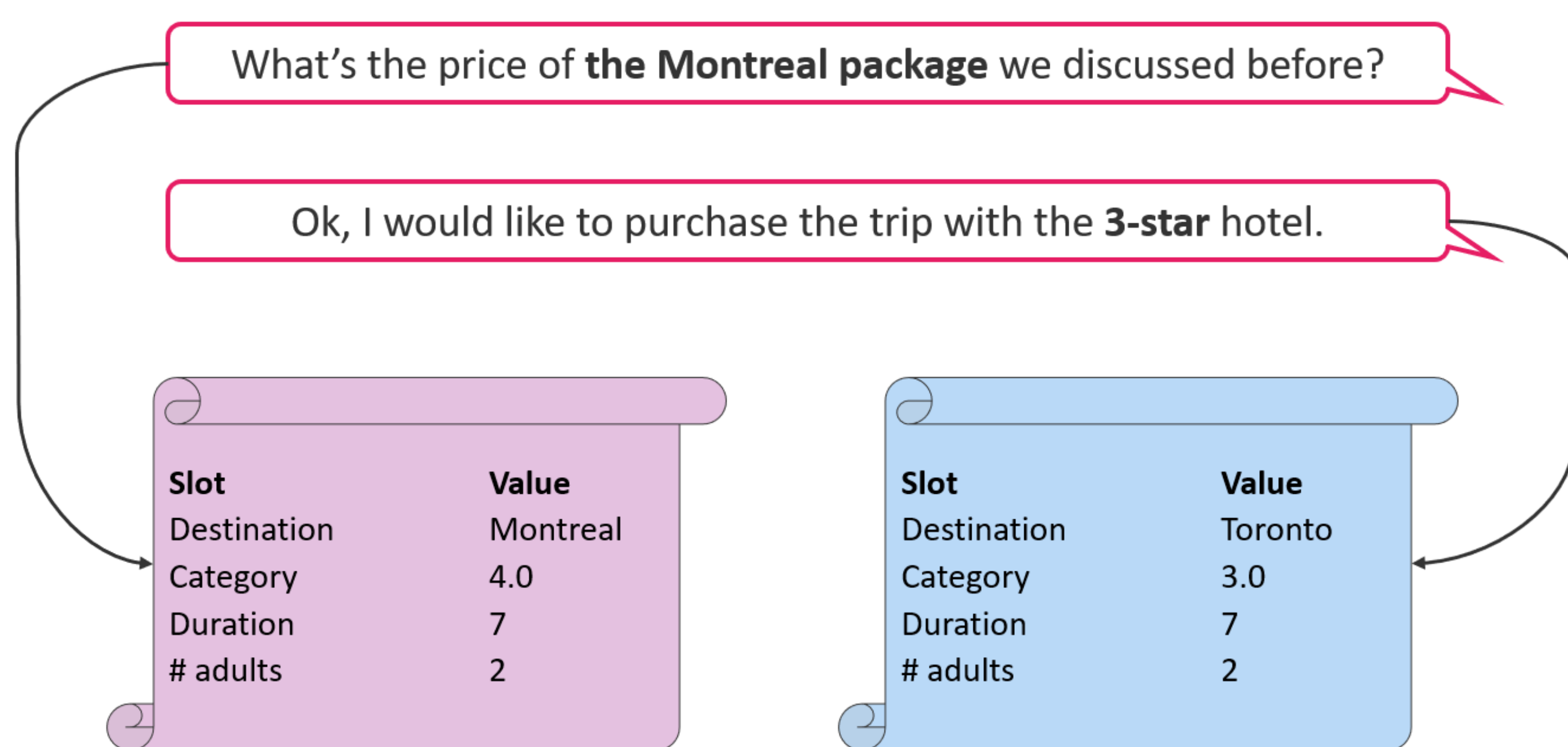


## Introduction

- This project focuses on frame tracking [1], a way to incorporate memory into a goal-oriented dialogue system.
- We propose a model with attention mechanism that is inspired by how human resolves frame references.
- We introduce a method to generate synthetic frame tracking data and use the data for pre-training.

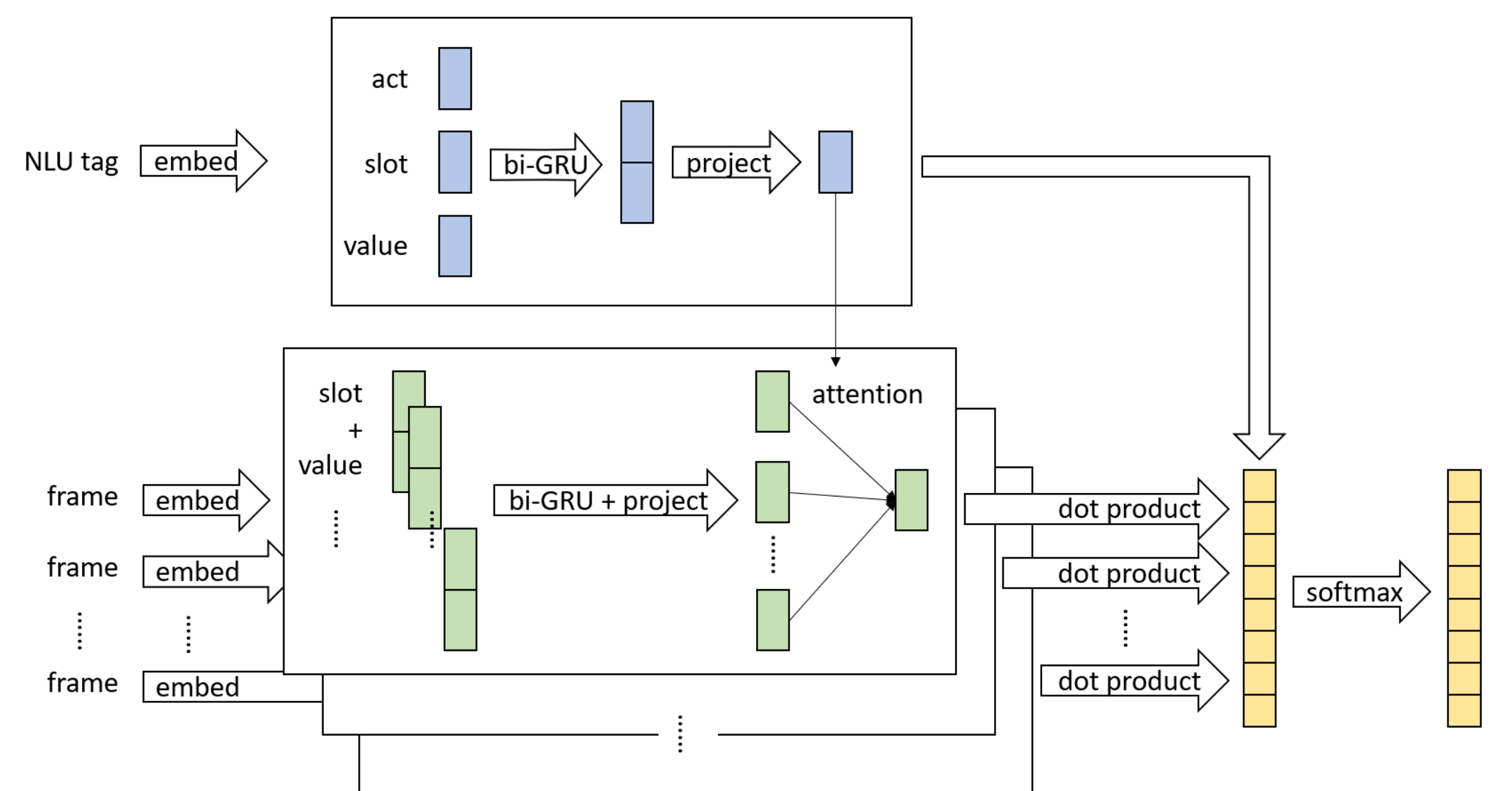
## Frame tracking

- In a goal-oriented dialogue, a frame summarizes an option for the user. It is represented as a list of slot-value pairs.
- The goal of frame tracking is to find out the most related frame for each natural language tag in the utterance.
- We use FRAMES dataset [1] as the training set.



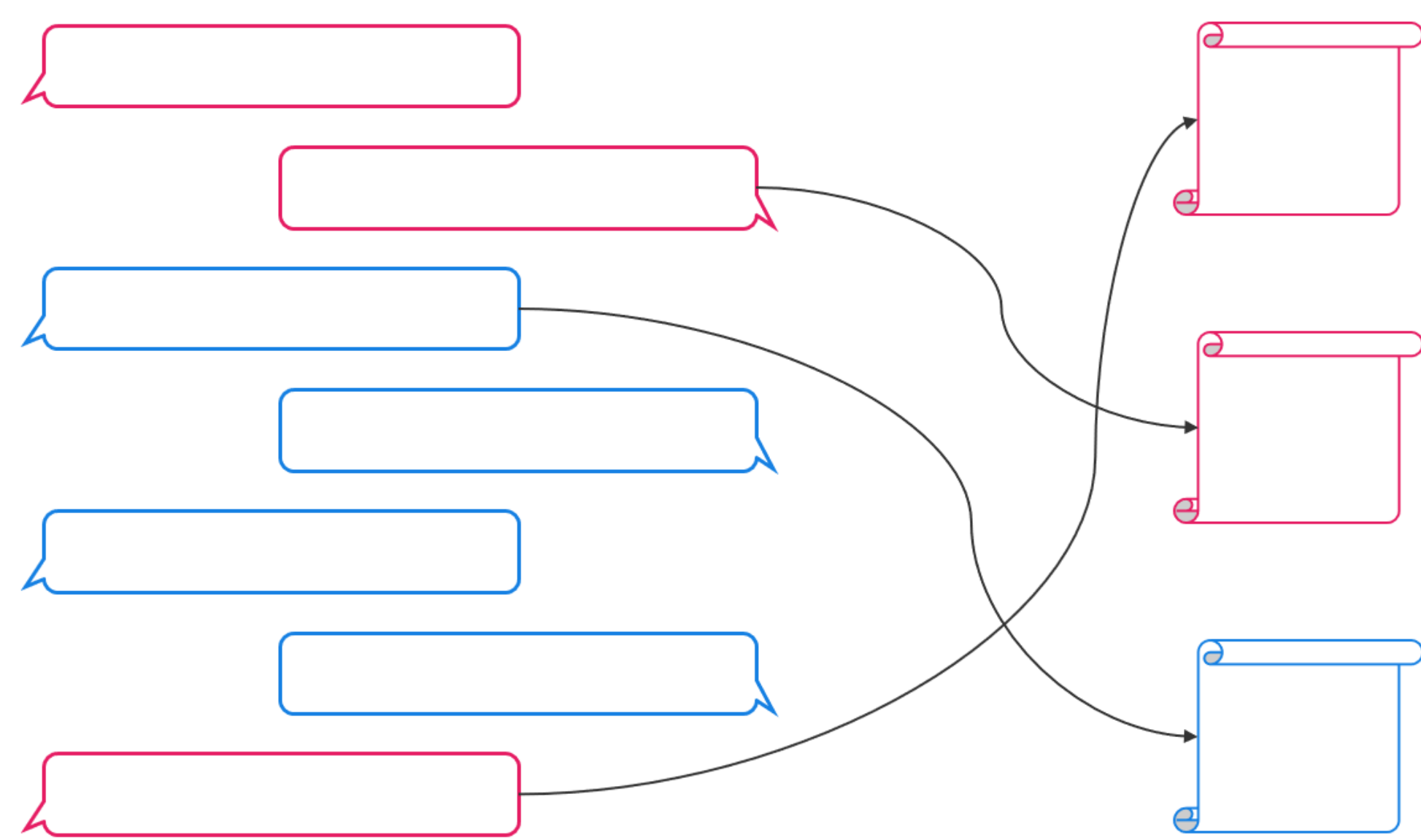
## Model

- We use a lookup table for act and slot embedding.
- We consider two types of text embedding for value tags: pre-trained BERT feature and GRU-based embedding training from scratch.
- We experiment with several attention mechanisms, including dot product, query-key attention, etc.



## Synthetic datasets

- We generate synthetic frame tracking data by interleaving dialogues from MultiWOZ, a large dataset without frame tracking labels.
- We transform dialogue state labels into frames, and create synthetic frame references.

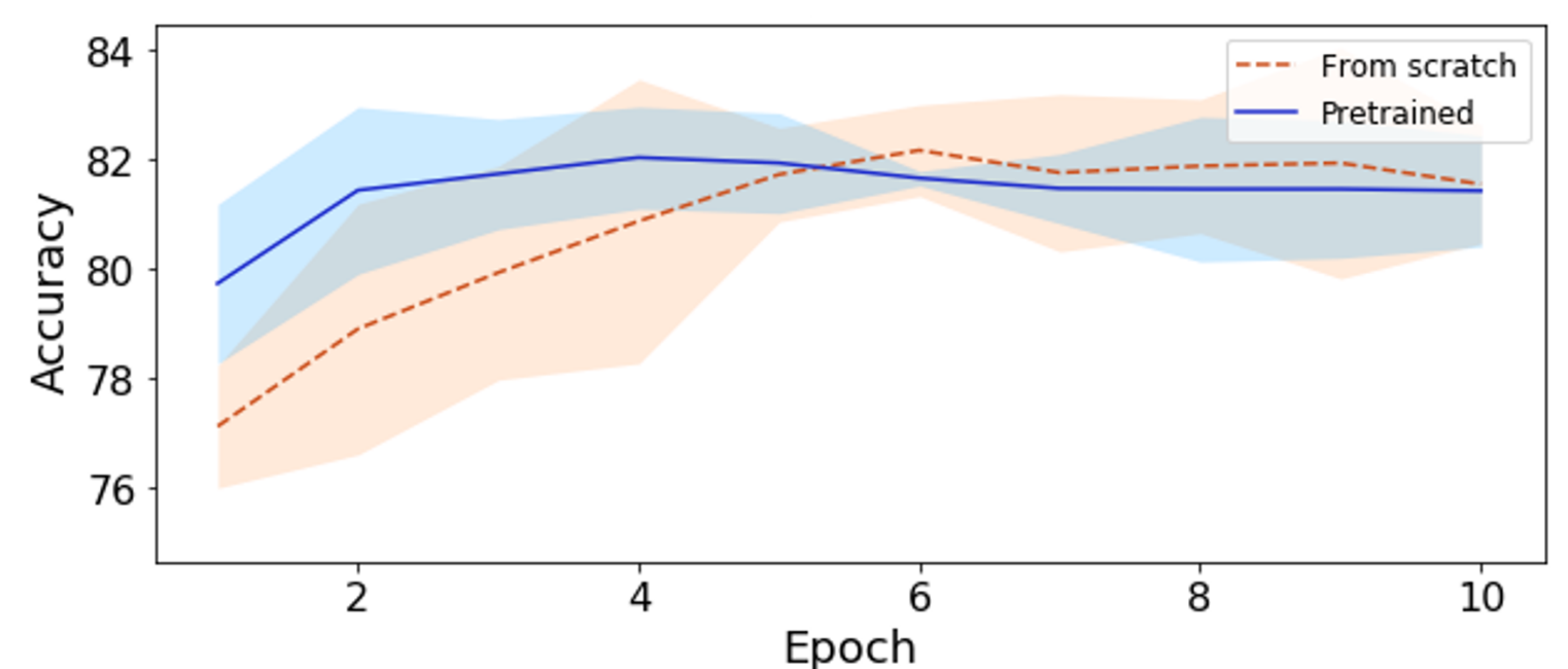


Dataset	# dialogues	Domain(s)
FRAMES	1369	Hotel + flight
Synthetic 1	5488	Single domain
Synthetic 2	3820	Hotel + restaurant
Synthetic 3	3820	Hotel + transportation (taxi and train)

## Results

- We evaluate the model by the accuracy of frame reference prediction on FRAMES dataset.
- The model achieves 83.1% accuracy after hyperparameter tuning.

Methods	No attention	Attention (dot product)
Random	58.1 ± 0.22	-
Maluuba [2]	76.4 ± 4.49	-
BERT w/o shuffle	77.5 ± 0.52	81.0 ± 0.69
GRU w/o shuffle	79.3 ± 0.28	81.9 ± 1.05
BERT w/ shuffle	81.0 ± 0.73	82.3 ± 1.70
GRU w/ shuffle	79.5 ± 0.65	<b>82.8 ± 0.52</b>



## Conclusion

- We propose a frame tracking model with attention mechanism, and improve the accuracy by 6.7 percentage point.
- The model pre-trained on synthetic frame tracking data converges faster comparing to the model training from scratch.

## References

- [1] Layla El Asri, Hannes Schulz, Shikhar Sharma, Jeremie Zumer, Justin Harris, Emery Fine, Rahul Mehrotra, and Kaheer Suleman. Frames: A corpus for adding memory to goal-oriented dialogue systems. *arXiv preprint arXiv:1704.00057*, 2017.
- [2] Hannes Schulz, Jeremie Zumer, Layla El Asri, and Shikhar Sharma. A frame tracking model for memory-enhanced dialogue systems. *arXiv preprint arXiv:1706.01690*, 2017.