Question Answering for a Football/Soccer bot

London Football Club
Nikhil Doifode, Neeraj Ganu

Problem Statement

Objective: Create multi-hop Question Answering (QA) system for non-goal oriented soccer related conversations

- Creating RNN-Attention based system for QA
- Integrating knowledge graph for adding domain knowledge

Give an example for the task

Using a KG-Copy Network for Non-Goal Oriented Dialogues

USER:	Hey what's up?
SYSTEM:	Nothing much, how are you? Do you want to talk soccer?
USER:	I am fine, thanks. Yeah, do you know what is the home ground of Arsenal ?
SYSTEM:	Arsenal's home ground is *Emirates Stadium*.
USER:	Oh ok, what's its capacity?
SYSTEM:	It has a capacity of *60,338*.

Subject	Predicate	Object
Arsenal	Home venue	Emirates Stadium
Emirates Stadium	capacity	60,338
Arsenal	chairman	Chips Keswick
Arsenal	Head coach	Unai Emery

Arsenal Knowledge Graph snippet.

Source: https://arxiv.org/pdf/1910.07834.pdf

What are the main challenges in this task?

- Creating QA for domain specific non-goal oriented conversations is hard
- Generating well articulated responses which are grounded in knowledge while keeping context
- Keeping Knowledge Graph updated and using it for fact based questions

How is this task typically addressed?

- QA for domain specific non-goal oriented conversations and generating well articulated responses is addressed by using hierarchical RNN based architecture models
- Keeping the responses grounded in truth is achieved by integrating Knowledge Graphs in the response generation (Eric, M. et al.: Key-value retrieval networks for task-oriented dialogue)
- Updating the Knowledge graphs - GUpdater

What is the specific system you are using?

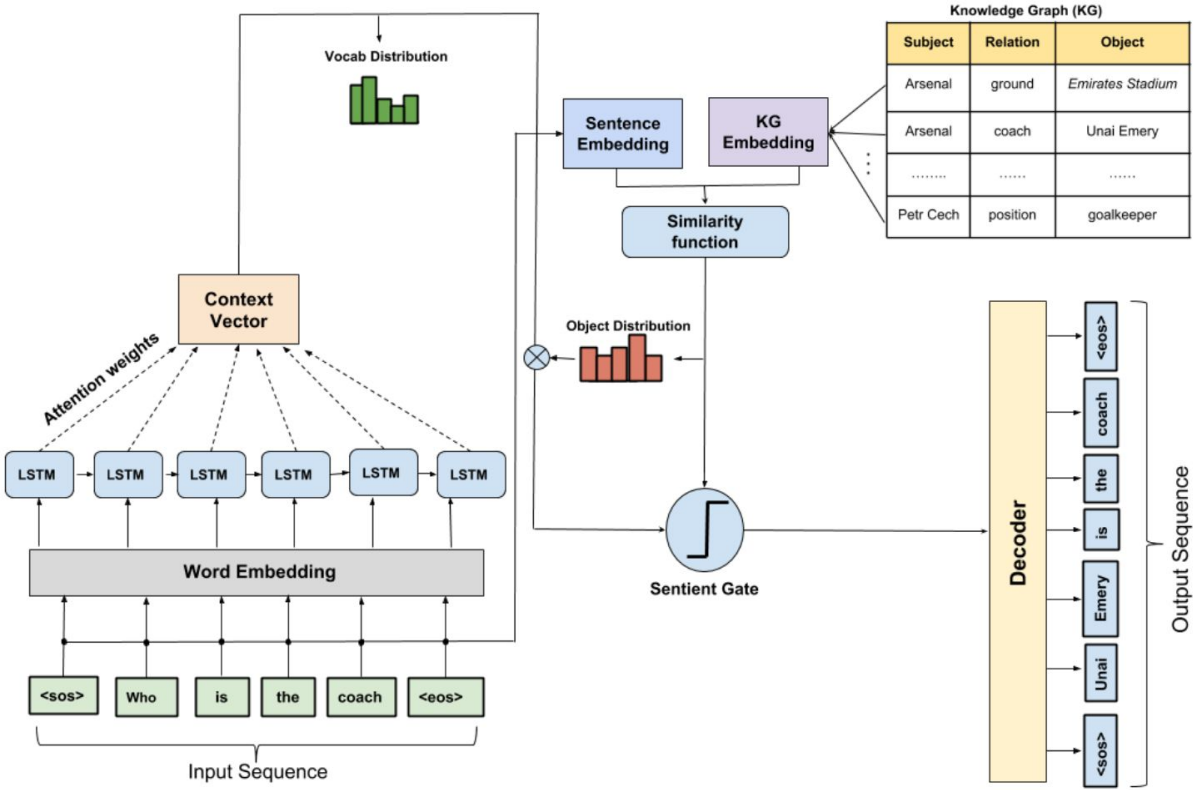


Fig. 4. KG-Copy Model Encoder-Decoder Architecture for Knowledge Grounded Response Generation.

Source: https://arxiv.org/pdf/1910.07834.pdf

What is the specific issue you are addressing?

- Knowledge Graph Updation
- Deciding between Knowledge graph output and RNN output for response generation in sentient gate

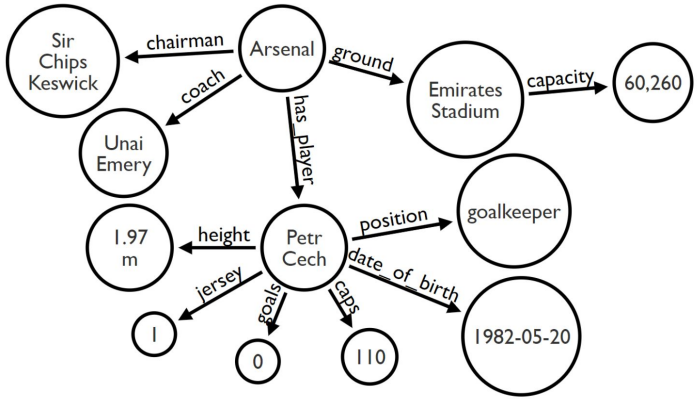
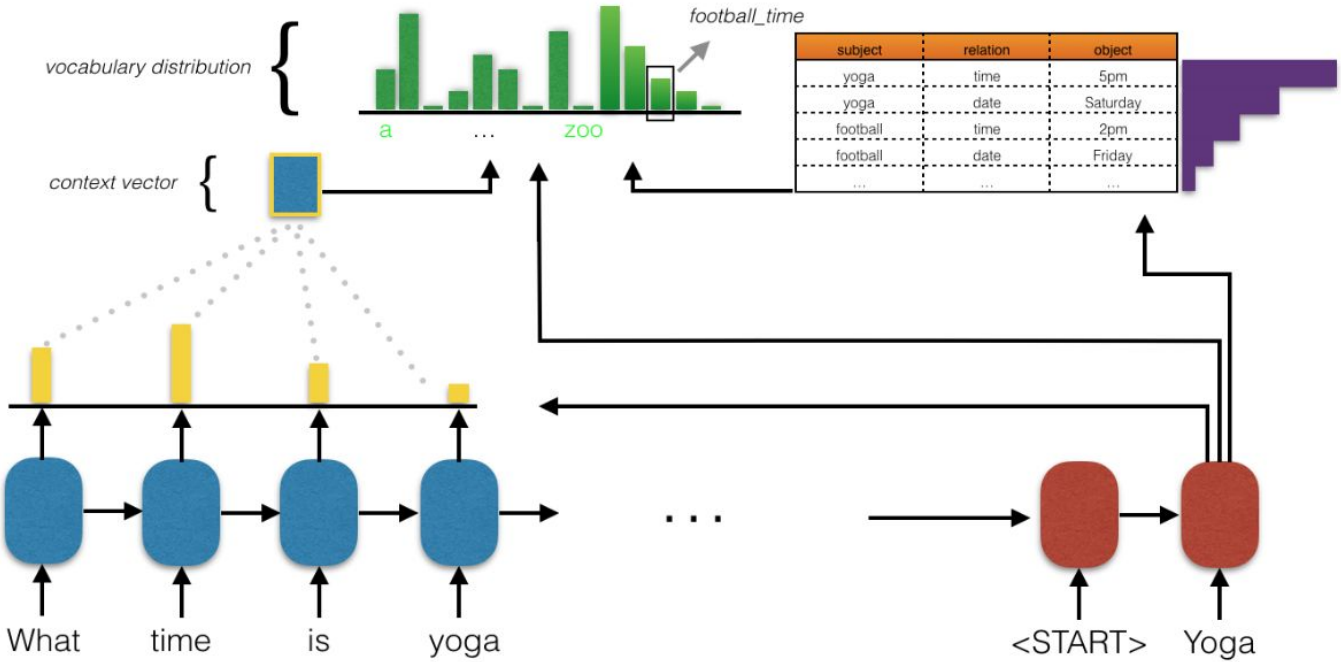


Fig. 3. Schema of the proposed Knowledge Graph for Arsenal.

How are you addressing this challenge?

- Updating knowledge graph
  - GUpdater like approach - too complex, needs data in specific format
  - Web scraping - we can get relations the way we want from wikipedia based scraper
- Sentient Gate choosing to ignore KG
  - Making changes in model
  - Improve KG similarity score for better consideration

Model from: Eric, M., Krishnan, L., Charette, F., Manning, C.D.: Key-value retrieval networks for task-oriented dialogue https://arxiv.org/pdf/1909.07606v1.pdf



Initial Results

Statistics	Count
Total Vocabulary Words (v)	4782
Avg. Number of Conversations/team	83
Avg. Number of Triples/team	148
Avg. Number of Entities/ team	108
Avg. Number of Relations/team	13

Dataset Statistics

Model	BLEU		Entity-F1	
	Valid	Test	Valid	Test
Vanilla Encoder-decoder with Attention	1.04	0.82	_	_
Mem2Seq [20]	1.30	0.52	6.78	7.03
KG Copy (proposed model)	**2.56**	**2.05**	24.98	23.58

Reference Result

Model	BLEU		Entity-F1	
	Valid	Test	Valid	Test
Our KG Copy	2.43	2.0	20.40	20.87

Results with our changes

Initial Results

KG from paper's dataset

Subject	Predicate	Object
Arsenal	ground	Emirates Stadium
Emirates Stadium	capacity	60,704
Arsenal	chairman	Sir Chips Keswick
Arsenal	coach	Unai Emery

Our updated KG

Subject	Predicate	Object
Arsenal	ground	Emirates Stadium
Emirates Stadium	capacity	60,704
Arsenal	chairman	
Arsenal	coach	Mikel Arteta

Conclusions

- Domain specific knowledge graphs are helpful in getting better results for factoid QA systems
- Even though current model uses knowledge graph sparsely for specific task but it is still better than the models which do not use knowledge graphs (E.g. mem2seq)
- Recent work has shown that in domain-specific tasks (including finance, law, and medicine), K-BERT significantly outperforms BERT, which demonstrates that K-BERT is an excellent choice for solving the knowledge-driven problems that require experts

Thank you!

Questions?