On Incorporating Structural Information to Improve Dialogue Response Generation

Nikita Moghe[†], Priyesh Vijayan[⇑], Balaraman Ravindran[*+], Mitesh M. Khapra[*+]

*Indian Institute of Technology Madras

+ Robert Bosch Centre for Data Science and Artificial Intelligence, IIT Madras

† University of Edinburgh ⇑ McGill University & MILA

{cs16s016, priyesh, ravi, miteshk}@cse.iitm.ac.in

Motivation

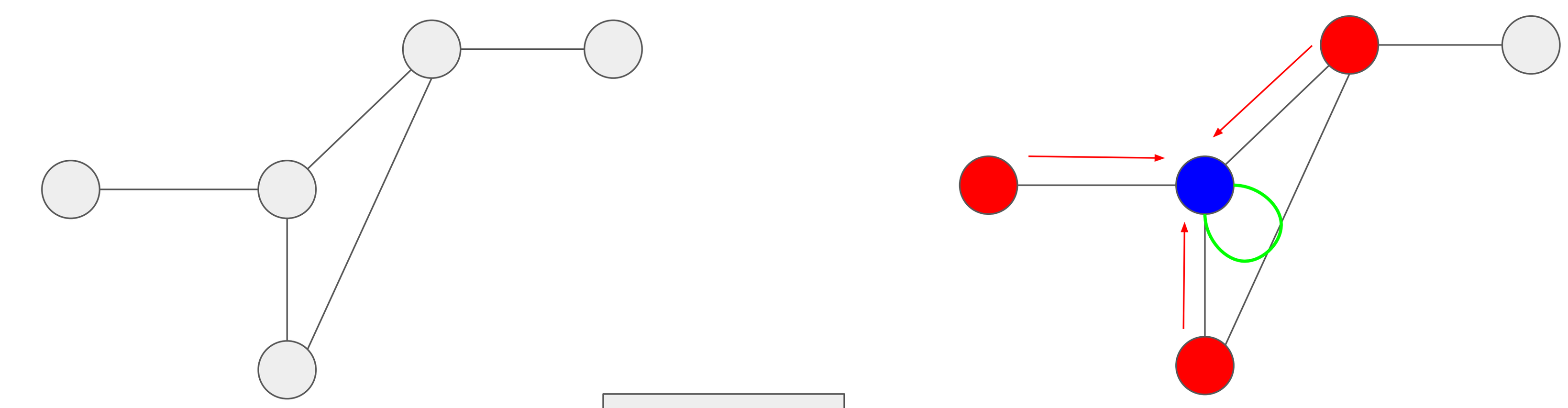
Structural information is ubiquitous in natural language

Dependency parses, constituency trees, co-reference graphs

But deep neural methods hope to learn these linguistic properties implicitly from data during training

Can we *explicitly* incorporate structural information in these neural architectures?

Graph Convolutional Network (GCN)



Undirected Graph

Node information

$$h_i = ReLU\left(W_0 h_i + \sum_{u \in N(i)} W_1 h_u\right)$$

Neighbourhood information

Background Aware Conversation Systems

Task: Given conversation history and associated background knowledge, generate a response

Dataset: Holl-E Domain: Movies Language: English Stats: ~9k chats , ~90k utterances, ~9k resources

Every alternate response is formed by copying words from the resource with appropriate prefixes/ suffixes

Speaker 1: Yes very true, this is a real rags to riches story. Russell Crowe was excellent as usual

Speaker 2: Russell Crowe owns the character of James Bradock. He's a good fighter turned hack. Injury, bad luck and the Depression sends him down the drain.

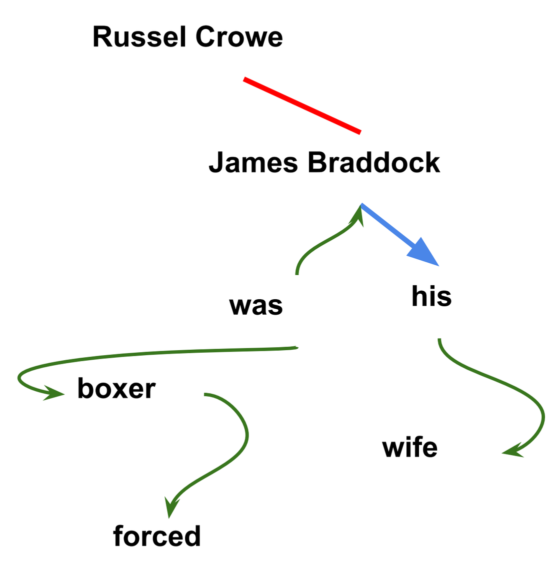
Speaker 1: Totally! Oh by the way do you remember his wife ... how she wished he would stop

Speaker 2: Yes! **His wife Mae had prayed for years that he would quit boxing, before becoming permanently injured.**

Conversation

Russell Crowe plays the lead. At this point James Braddock was a boxer, who was forced to retired from the ring after breaking his hand in his last fight. His wife Mae had prayed for years that he would quit boxing, before becoming permanently injured. To support his family, Braddock works as a laborer at the docks, but he still has a dream to box.

Resource

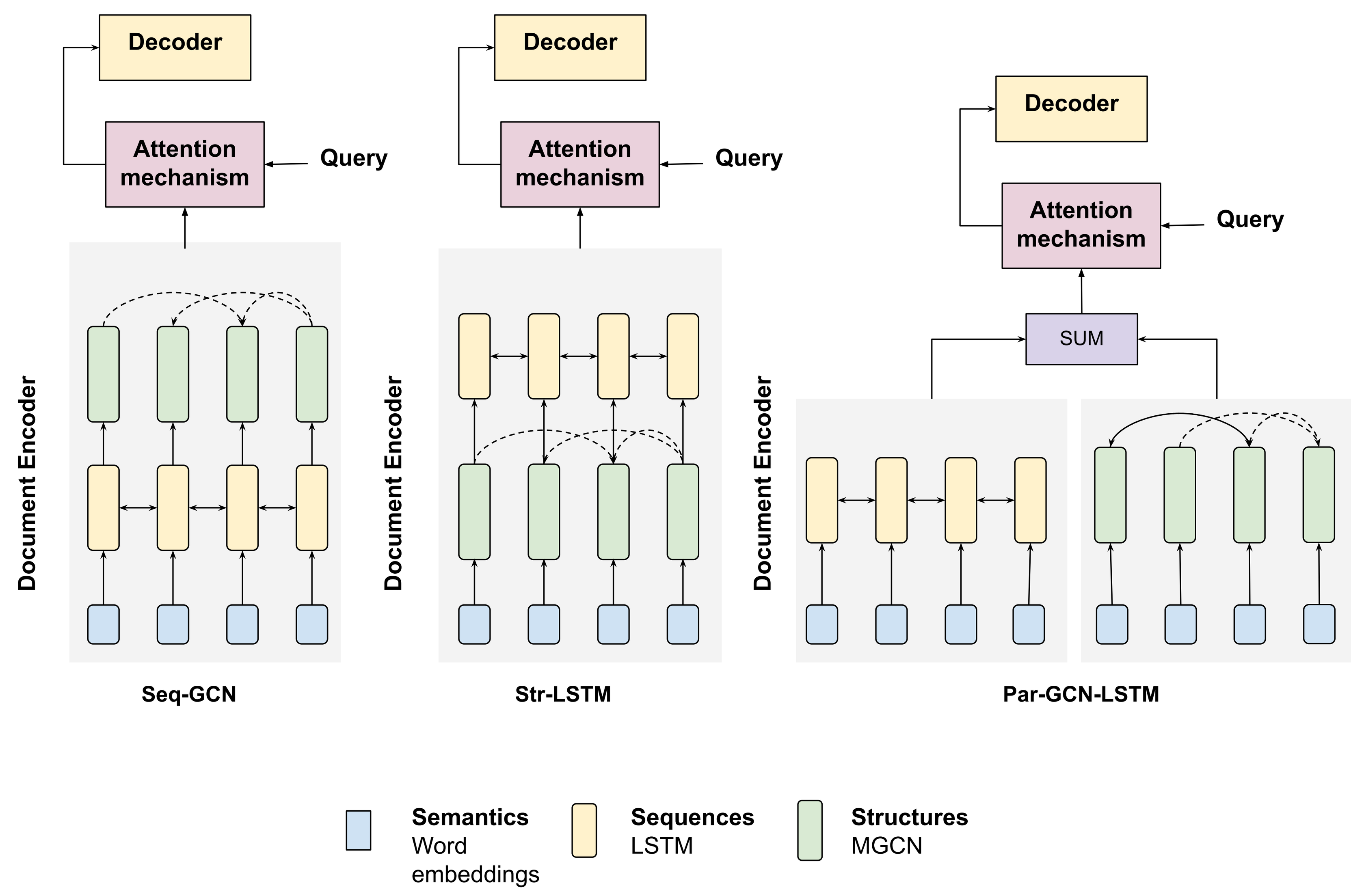


Dependency graph

Entity graph

Co-reference graph

Semantics-Sequences-Structures Framework



Multi Graph GCN: $h^{k+1}_i = ReLU\left(W^k_0 h^k_i + \sum_{N \in G} \sum_{u \in N(i)} \left(W^k_{dir(i,u)} h^k_u + b^k_{L(i,u)}\right)\right)$

Results

	Model	BLEU	ROUGE		
			1	2	L
	HRED	05.23	24.55	07.61	18.87
	GTTP	13.92	30.32	17.78	25.67
	BiDAF	16.79	26.73	18.82	23.58
GloVe	Sem	04.40	29.72	11.72	22.99
	Sem + Seq	14.83	36.17	24.84	31.07
	Sem + Seq + Str	18.96	38.61	26.92	33.77
ELMo	Sem	14.36	32.04	18.75	26.71
	Sem + Seq	14.61	35.54	24.58	30.71
	Sem + Seq + Str	19.32	39.65	27.37	34.86
BERT	Sem	11.26	33.86	16.73	26.44
	Sem + Seq	18.49	37.85	25.32	32.58
	Sem + Seq + Str	**22.78**	**40.09**	**27.83**	**35.20**

Analysis

Conclusion and Future Work

- We demonstrate the usefulness of incorporating structural information into the standard semantics+sequential neural models.
-
- In future, we can design architectures that have the power of both Recurrent Neural Networks and Graph Convolutional Networks

References

[1] *Semi-supervised classification with Graph Convolutional Network,* Thomas N Kipf and Max Welling., In International Conference on Learning Representations (ICLR) 2016

[2] *Towards Exploiting Background Knowledge for Building Conversation Systems*, Nikita Moghe, Siddhartha Arora, Suman Banerjee, and Mitesh M. Khapra. In proceedings of Empirical Methods in Natural Language Processing (EMNLP) 2018

This work was done by Nikita and Priyesh at IIT Madras