An Empirical Study of Graph Neural Networks Based Semantic Parsing

Fengyuan Xu

Nanjing Univ.

Joint work with Shucheng Li, Lingfei Wu, Shiwei Feng, Fangli Xu, and Sheng Zhong.

What is semantic parsing?

Are there any jobs with Microsoft involving sql ?

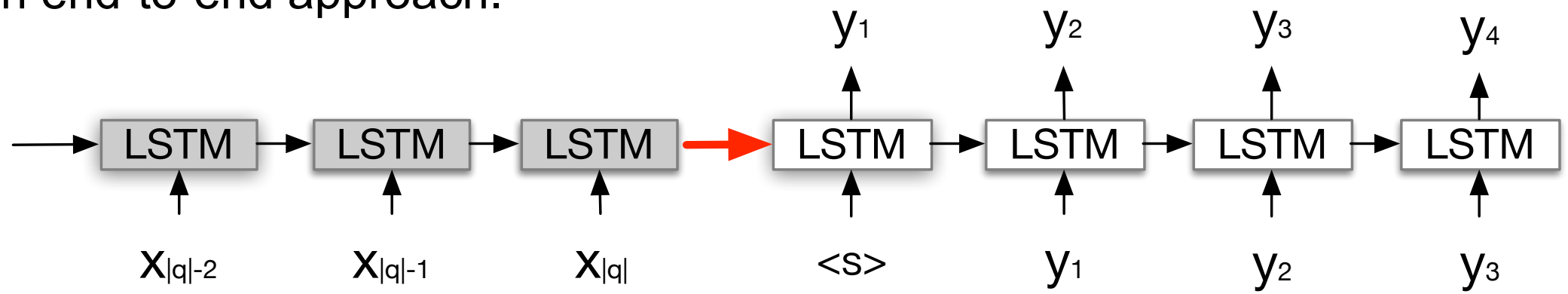
After semantic parsing

Answer(A,(job(A),company(A,C),const(C,'Microsoft'),language(A, L),const(L,'sql'))))

Previous semantic parser

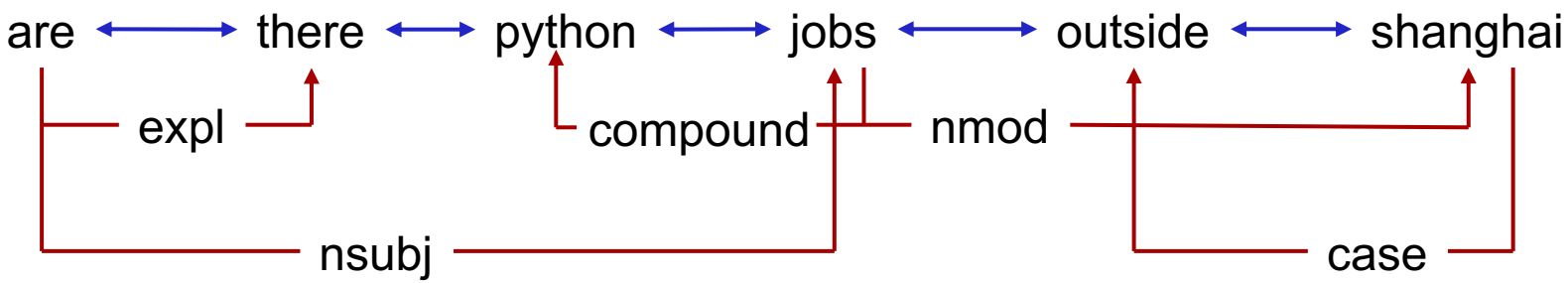
Translate problem: from natural language sentences to hierarchical output like logical forms

- Seq2seq model: an encoder-to-decoder architecture, an end-to-end approach.



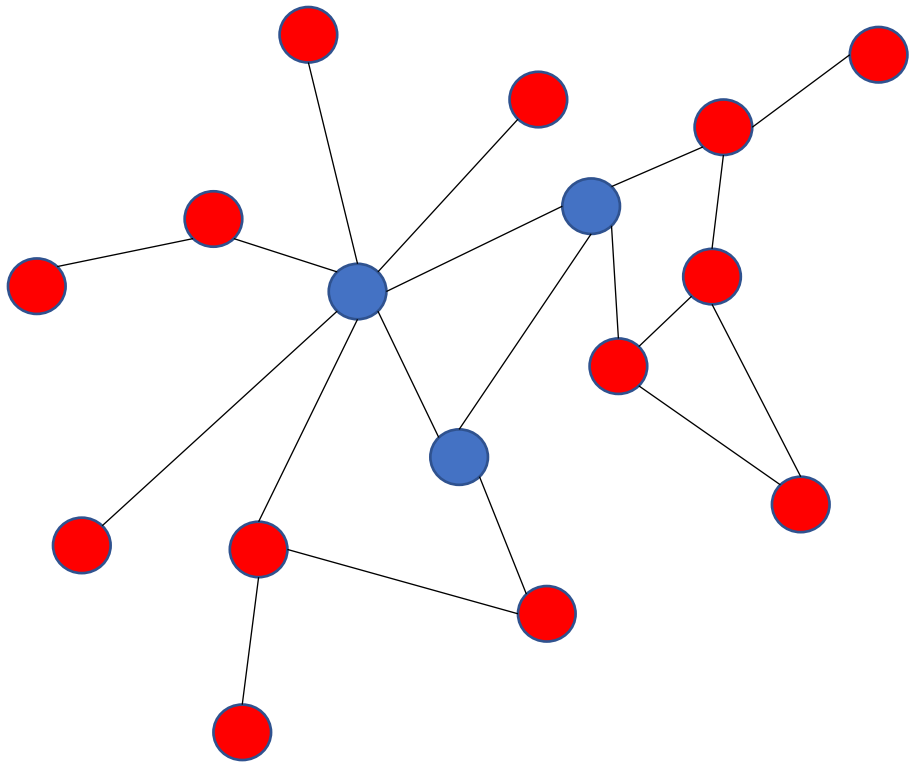
Problem

- Sequence encoder and decoder encode word order features but ignore more syntactic information such as dependency and constituency features.



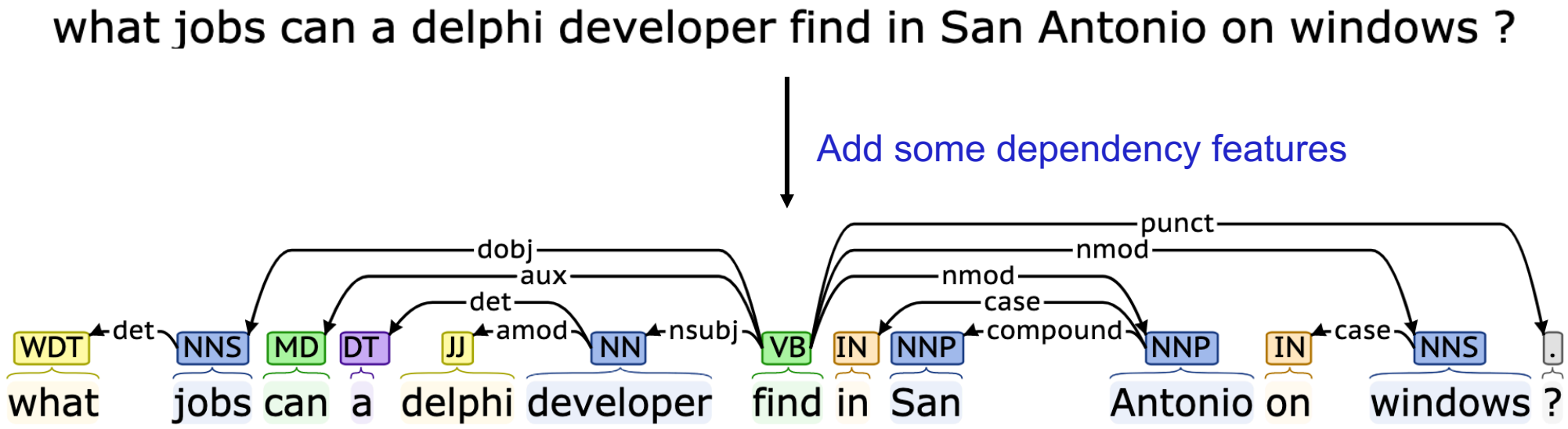
Why should we choose Graph Neural Network(GNN) for semantic parsing task?

- Some inputs can be naturally expressed in graph, for example, protein network



Why should we choose Graph Neural Network(GNN) for semantic parsing task?

- Hybrid Graph with some augmented information

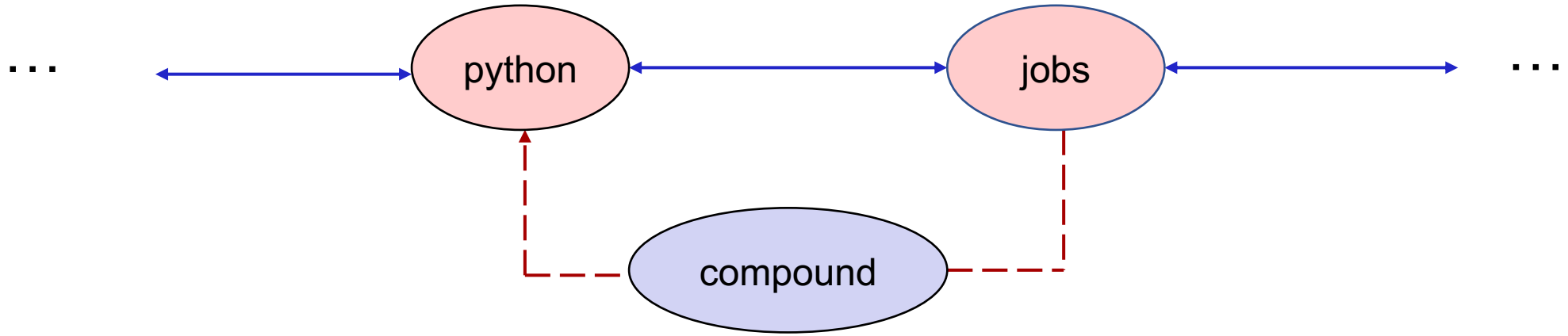


Key word "jobs" is far away from its attribute "windows", but with dependency features, it's more easy to learn a high-quality input representation (from 10 hops to 2 hops in graph).

Our approach

- How to construct our graph input ?

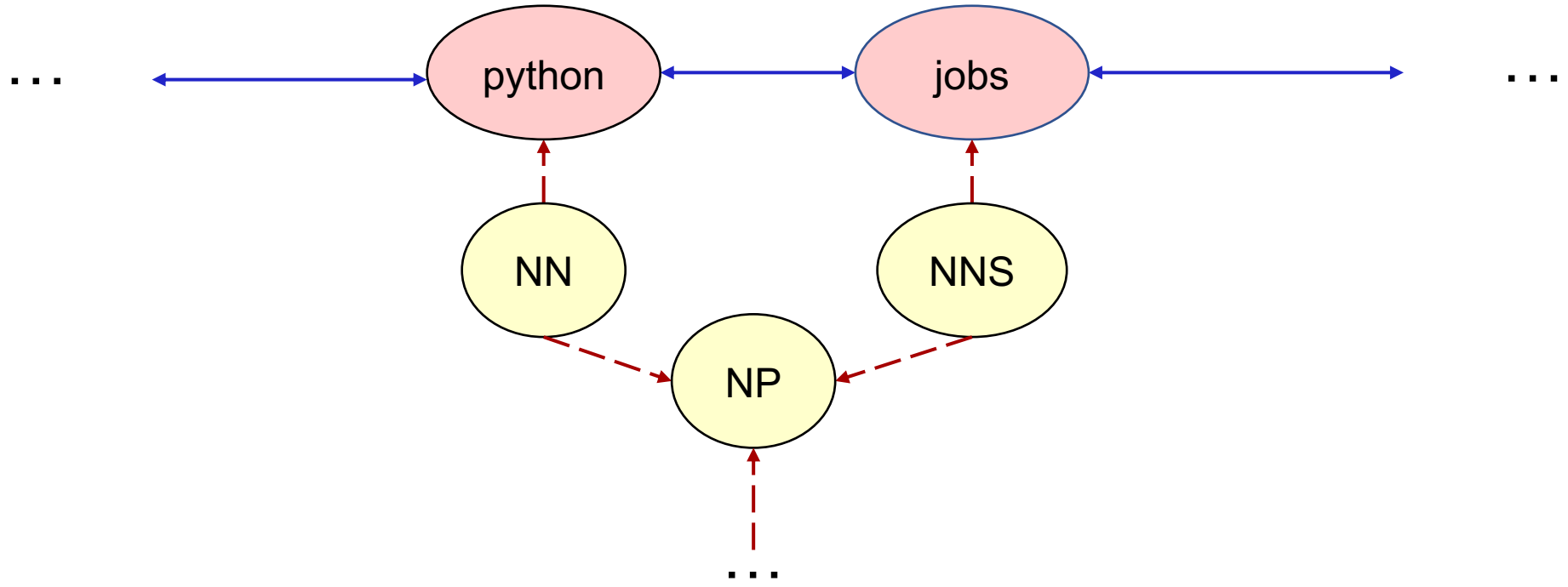
1) dependency parsing tree, turn dependency features to new nodes and link them with two nodes in sentence.



Our approach

- How to construct our graph input ?

2) constituency parsing tree, incorporate constituency tree nodes and word nodes into a single graph.



Our approach

- How to learn good representation from this kind of graph and get high-quality hierarchical output ?

1) Graph encoder

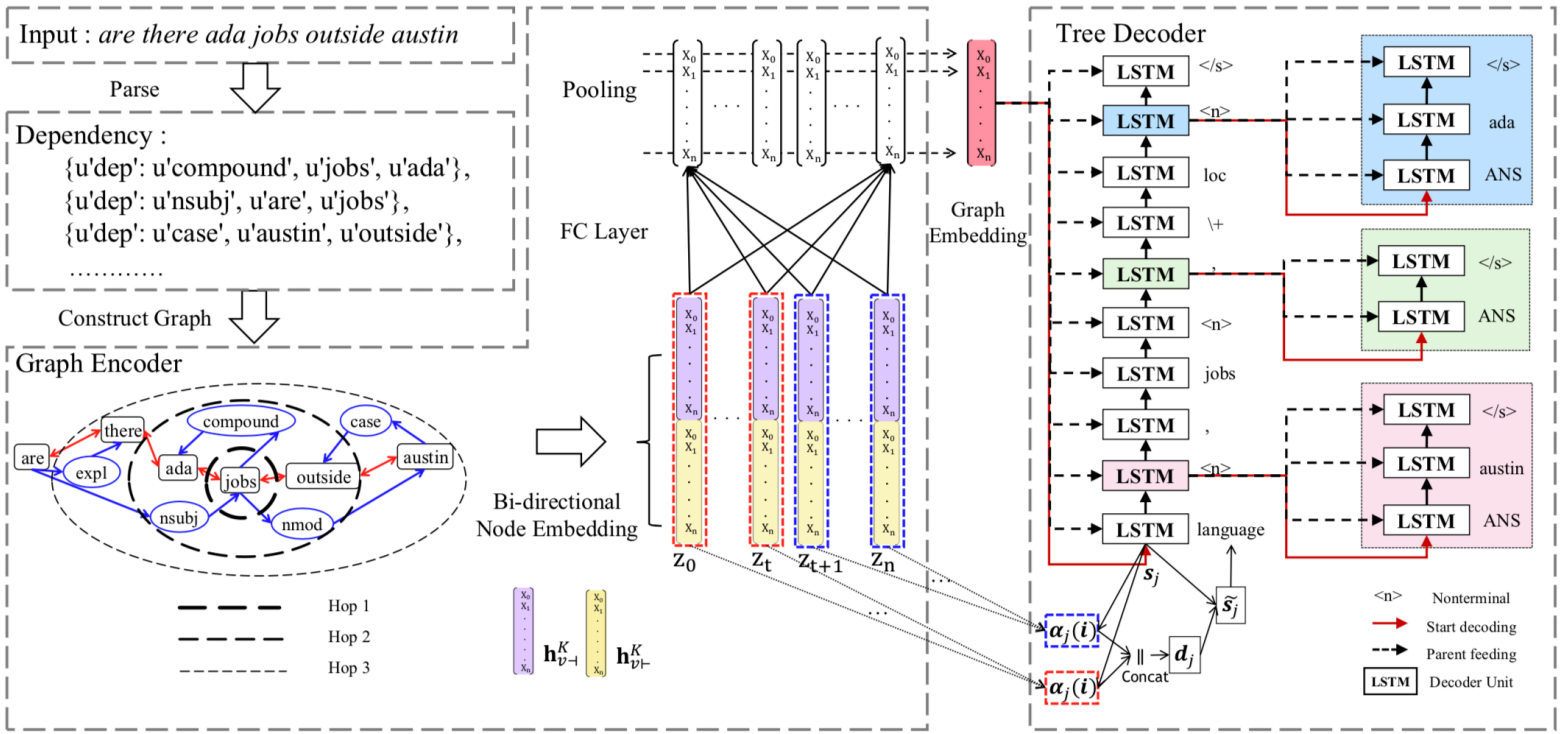
- Our graph encoder adapts an inductive node embedding algorithm. We generate each node embedding in graph by bi-directionally aggregating information from its local neighborhood in *K* hops.

2) Hierarchical decoder

- Logic form output are usually hierarchical, and traditional sequence decoder need to memory some auxiliary information like brackets to keep the structure.
- So we adapt a tree decoder in our model, it decodes logic form outputs in a coarse-to-fine way.

Our approach: Graph2Tree

- Use Bi-LSTM to process word order features
- Calculate different attention vectors for different types of nodes in graph input



Experiment

- Datasets
  - Jobs640
  - Geo880
  - ATIS

- Baseline
  - Graph2Seq model consists of graph encoder and sequence decoder[1]
  - Seq2Seq model and Seq2Tree[2] model

[1] Kun Xu*, Lingfei Wu*, Zhiguo Wang, Yansong Feng, Michael Witbrock, and Vadim Sheinin (both authors contributed equally), "Graph2Seq: Graph to Sequence Learning with Attention-based Neural Networks", arXiv preprint 2018.
[2] Li Dong and Mirella Lapata. 2016. Language to logical form with neural attention. arXiv preprint arXiv:1601.01280 (2016).

Results

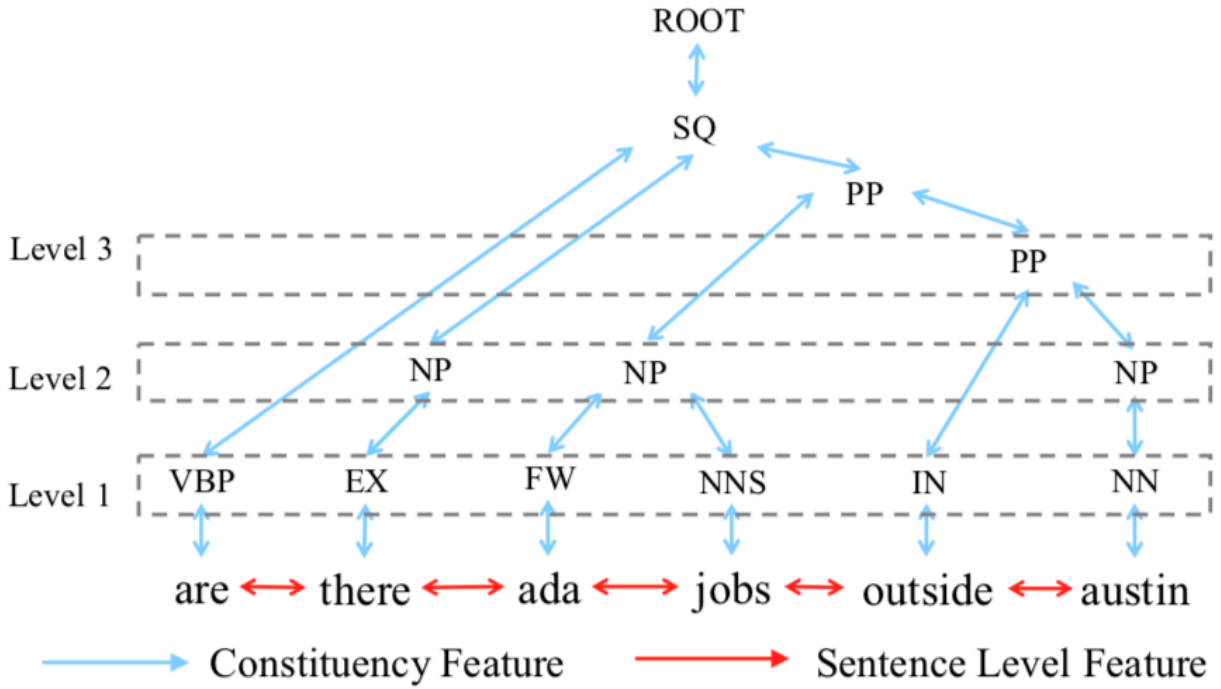
Methods	*JOBS*	*GEO*	*ATIS*
Jia and Liang (2016)	-	85.0	76.3
Dong and Lapata (2016)-Seq2Seq	87.1	85.0	84.2
Dong and Lapata (2016)-Seq2Tree	90.0	87.1	84.6
Rabinovich et al. (2017)	92.9	85.7	**85.3**
Xu and Wu (2018)-Graph2Seq	88.6	85.7	83.8
Graph2Tree	**92.9**	**88.2**	84.6

- Graph2tree model achieves remarkable performance in JOBS & GEO
- In bigger datasets like ATIS, Graph2tree has limited performance

Results

- Syntactic graph with constituency tree features from a simple and short sentence

"Are there ada jobs outside austin?"



- Syntactic graph with some noise or complex structure may be detrimental to semantic parsing.

Results

- Try constituency parsing tree with different structure complexity.
- Remove some noise nodes in dependency parsing tree manually.

Methods	Graph2Tree	Graph2Seq
Word Order	91.4	87.1
Output correction: WO+DEP	92.1	88.6
Original condition: WO+DEP	91.5	85.0
Constituency two layer cut: WO+CON	91.4	86.4
Constituency one layer cut: WO+CON	**92.9**	88.6
Constituency tree all layer: WO+CON	92.1	85.0

Why Graph2tree?

- Decoding results compared to baseline

Methods	Prediction results
Reference str	job (ANS) , language (ANS , 'delphi') , title (ANS , 'developer') , loc (ANS , 'san antonio') , platform (ANS , 'windows')
Graph2tree	job (ANS) , language (ANS , 'delphi') , title (ANS , 'developer') , loc (ANS , 'san antonio') , platform (ANS , 'windows')
Graph2seq	job (ANS) , language (ANS , 'delphi') , title (ANS , 'developer') , platform (ANS , 'windows')
Seq2seq	job (ANS) , language (ANS , 'delphi') , title (ANS , 'developer') , loc (ANS , 'san antonio')

Conclusion

- We presented Graph2Tree model, a generalized model for graph inputs and tree structured outputs.

- We made some primary attempts on how to construct a proper choice of graph for this type of neural semantic parsing.

Thank You!

An Empirical Study of Graph Neural Networks Based Semantic Parsing

Joint work with Shucheng Li, Lingfei Wu, Shiwei Feng, Fangli Xu, and Sheng Zhong.