Semantic Parsing for Conversational Question Answering over Knowledge Graphs

Laura Perez-Beltrachini[1], Parag Jain[1], Emilio Monti[2], Mirella Lapata[1]

[1] School of Informatics, University of Edinburgh

[2] Amazon Alexa
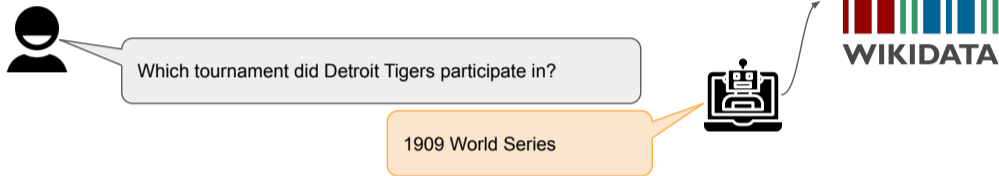
{lperez,parag.jain,mlap}@inf.ed.ac.uk, monti@amazon.co.uk

Conversational Question Answering over Knowledge Graphs

Conversational Question Answering over Knowledge Graphs

Conversational Question Answering over Knowledge Graphs

Conversational Question Answering over Knowledge Graphs

Grounding User Questions into SPARQL Queries over KGs

Which sports team was the champion of that tournament?

```
SELECT ?x WHERE { wd:          wdt:          ?x. ?x wdt:P31 wd:          . }
```

Grounding User Questions into SPARQL Queries over KGs

Which sports team was the champion of that tournament?

Q12973014   P1346   Q846847

```
SELECT ?x WHERE { wd:        wdt:        ?x. ?x wdt:P31 wd:        . }
```

Grounding User Questions into SPARQL Queries over KGs

Grounding User Questions into SPARQL Queries over KGs

Challenges

- Scale (thousands of types and relations, millions of entities)
- Conversation phenomena (ellipsis, coreference, clarifications)

Datasets for Semantic Parsing

	Open Domain	Conversational	Executable
ATIS, CoSQL, SParC [Suhr et al. (2018); Yu et al. (2019b,a)]		✓	✓
LC-QuAD2.0, ComplexWebQuestions [Dubey et al. (2019); Talmor and Berant (2018)]	✓		✓
CSAQ, ConvQuestions [Saha et al. (2018); Christmann et al. (2019, 2022)]	✓	✓	

Datasets for Semantic Parsing

	Open Domain	Conversational	Executable
ATIS, CoSQL, SParC [Suhr et al. (2018); Yu et al. (2019b,a)]		✓	✓
LC-QuAD2.0, ComplexWebQuestions [Dubey et al. (2019); Talmor and Berant (2018)]	✓		✓
CSAQ, ConvQuestions [Saha et al. (2018); Christmann et al. (2019, 2022)]	✓	✓	
SPICE	✓	✓	✓

The SPICE Dataset

- 47 different user question intents

  Simple questions
  What are the countries of those sports teams?

  Reasoning questions
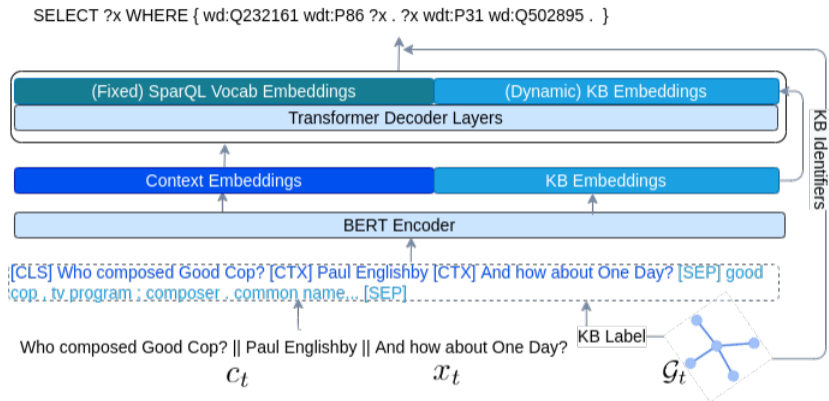 Which tournaments have less number of participating sports teams than 1909 World Series?

  Conversation
  And what about 1910 World Series?

- 197k instances

Avg. turn length	9.5
Avg. entities per conversation	7.6
Avg. types per conversations	6.5

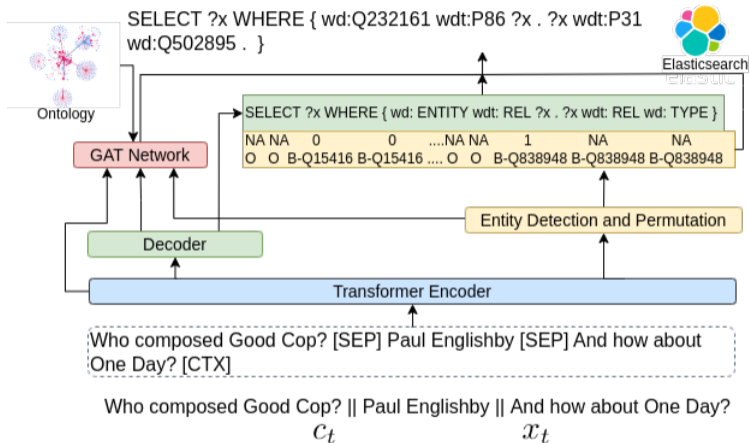# Sequence-to-sequence with Dynamic Vocabularies – **BertSP**

[Gu et al. (2021)]

- KB symbols as dynamic target vocabularies
- KB symbols from KB sub-graphs related to **mentioned entities**
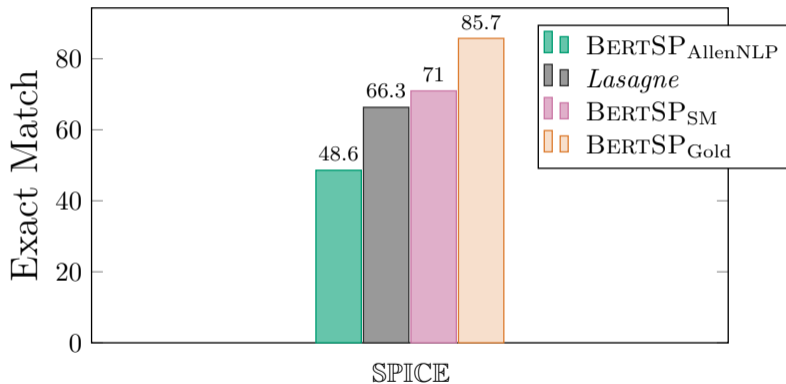- Previous turn as conversation context

Sparql Templates and Multiple Classifiers – **Lasagne**

[Kacupaj et al. (2021)]

- Different classifiers to predict KB symbols
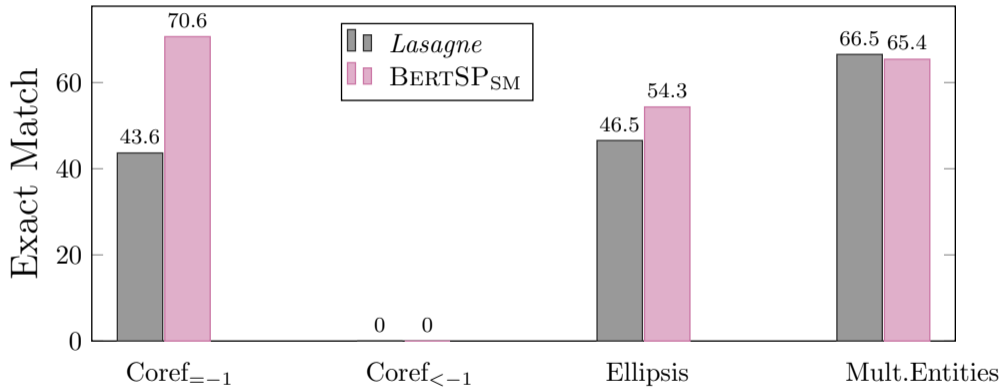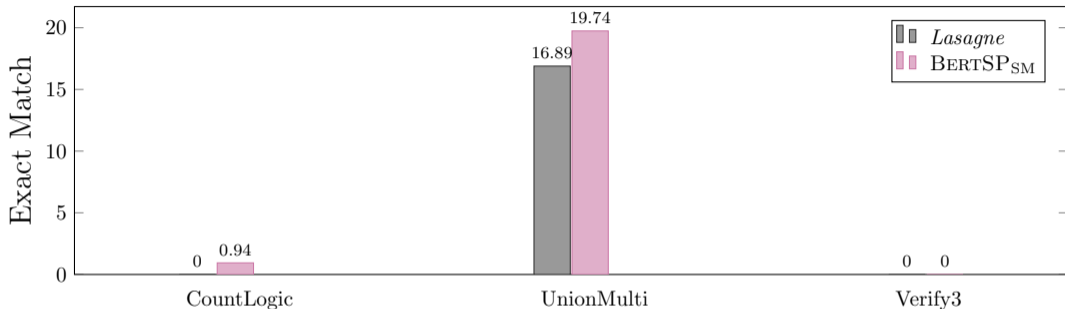- Previous turn as conversation context

Average Baseline Performance Over Question Types

Models Struggle on Specific Linguistic Phenomena

And to Generalise to Unseen Questions

- **CountLogic**: a count operation over a union operator
- **UnionMulti**: a union operator over two graph patterns with different relations
- **Verify3**: verification questions with three entities

Conclusions

Directions to improve conversational semantic parsing over large scale KGs:

- improve entity recognition and linking
- better modelling of conversation context
- models with better generalisation capabilities

Code and SPICE @ https://github.com/EdinburghNLP/SPICE

References I

Philipp Christmann, Rishiraj Saha Roy, Abdalghani Abujabal, Jyotsna Singh, and Gerhard Weikum. 2019. Look before you hop: Conversational question answering over knowledge graphs using judicious context expansion. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, pages 729–738, Beijing, China. ACM.

Philipp Christmann, Rishiraj Saha Roy, and Gerhard Weikum. 2022. Conversational question answering on heterogeneous sources. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 144–154, Madrid, Spain. ACM.

Mohnish Dubey, Debayan Banerjee, Abdelrahman Abdelkawi, and Jens Lehmann. 2019. Lc-quad 2.0: A large dataset for complex question answering over wikidata and dbpedia. In The Semantic Web – ISWC 2019, pages 69–78, Cham. Springer International Publishing.

Yu Gu, Sue Kase, Michelle Vanni, Brian Sadler, Percy Liang, Xifeng Yan, and Yu Su. 2021. Beyond i.i.d.: Three levels of generalization for question answering on knowledge bases. In Proceedings of the Web Conference 2021, WWW '21, page 34773488, New York, NY, USA. Association for Computing Machinery.

Endri Kacupaj, Joan Plepi, Kuldeep Singh, Harsh Thakkar, Jens Lehmann, and Maria Maleshkova. 2021. Conversational question answering over knowledge graphs with transformer and graph attention networks. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pages 850–862, Online. Association for Computational Linguistics.

Amrita Saha, Vardaan Pahuja, Mitesh Khapra, Karthik Sankaranarayanan, and Sarath Chandar. 2018. Complex sequential question answering: Towards learning to converse over linked question answer pairs with a knowledge graph. In Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), pages 705–713, New Orleans, Louisiana, USA. AAAI Press.

Alane Suhr, Srinivasan Iyer, and Yoav Artzi. 2018. Learning to map context-dependent sentences to executable formal queries. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 2238–2249, New Orleans, Louisiana. Association for Computational Linguistics.

Alon Talmor and Jonathan Berant. 2018. The web as a knowledge-base for answering complex questions. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers), pages 641–651, New Orleans, Louisiana. Association for Computational Linguistics.

Tao Yu, Rui Zhang, Heyang Er, Suyi Li, Eric Xue, Bo Pang, Xi Victoria Lin, Yi Chern Tan, Tianze Shi, Zihan Li, Youxuan Jiang, Michihiro Yasunaga, Sungrok Shim, Tao Chen, Alexander Fabbri, Zifan Li, Luyao Chen, Yuwen Zhang, Shreya Dixit, Vincent Zhang, Caiming Xiong, Richard Socher, Walter Lasecki, and Dragomir Radev. 2019a. CoSQL: A conversational text-to-SQL challenge towards cross-domain natural language interfaces to databases. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 1962–1979, Hong Kong, China. Association for Computational Linguistics.

References II

Tao Yu, Rui Zhang, Michihiro Yasunaga, Yi Chern Tan, Xi Victoria Lin, Suyi Li, Heyang Er, Irene Li, Bo Pang, Tao Chen, Emily Ji, Shreya Dixit, David Proctor, Sungrok Shim, Jonathan Kraft, Vincent Zhang, Caiming Xiong, Richard Socher, and Dragomir Radev. 2019b. SParC: Cross-domain semantic parsing in context. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 4511–4523, Florence, Italy. Association for Computational Linguistics.