IN5550: Neural Methods in Natural Language Processing
Lecture 12, part 2
Graph Neural Networks for NLP

Roxana Pop, Sondre Wold

University of Oslo



Contents



1. Graphs
2. Graph Neural Networks
3. GNNs in NLP
4. Question Answering
5. Pretraining language models
6. LLMs vs. KGs

Graphs

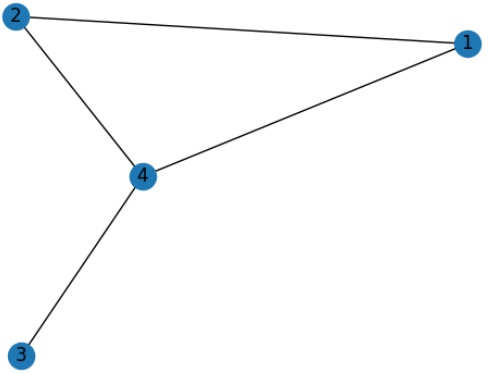
A graph G can be defined as a pair (V,E), where V is a set of vertices, and E is a set of edges between the vertices.

$V = \{1, 2, 3, 4\}$
$E = \{(1, 2), (2, 1), (2, 4), (3, 4), (4, 1), (4, 2), (4, 3)\}$

$G = (V, E)$

Neighborhood:

- $\mathcal{N}_4 = \{1, 2, 3\}$

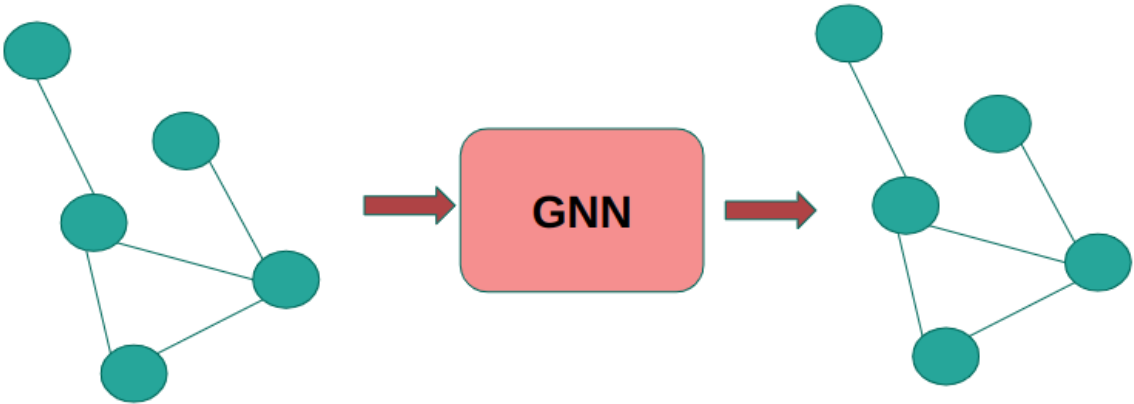


Contents



1. Graphs
2. Graph Neural Networks
3. GNNs in NLP
4. Question Answering
5. Pretraining language models
6. LLMs vs. KGs

GNNs - Graph Neural Networks



Contents



1 Graphs

2 Graph Neural Networks

3 GNNs in NLP

4 Question Answering

5 Pretraining language models

6 LLMs vs. KGs

GNNs in NLP

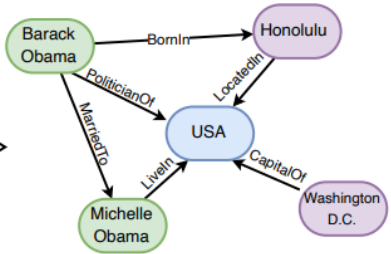
- A lot of applications
  - Huge survey by [Wu et al., 2021]
- Older methods – using GNNs to process a structured view of the sentence such as a dependency tree (eg. [Marcheggiani and Titov, 2017])
- Can use them to incorporate external knowledge in the form of knowledge graphs
  - Question answering
  - Pre-training language models
  - ...

Knowledge Graphs

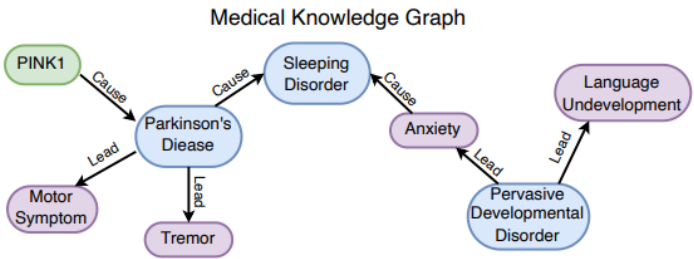
Encyclopedic Knowledge Graphs



Wikipedia



Domain-specific Knowledge Graphs



Taken from [Pan et al., 2023]

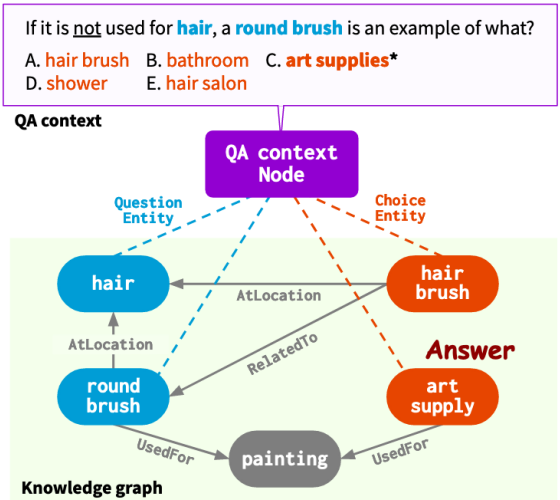
Contents



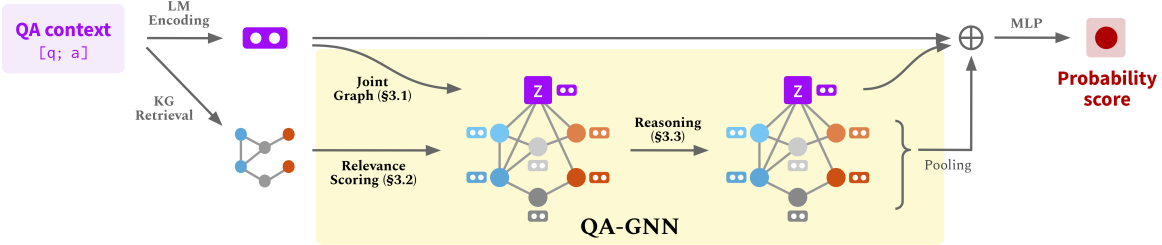
1. Graphs
2. Graph Neural Networks
3. GNNs in NLP
4. Question Answering
5. Pretraining language models
6. LLMs vs. KGs

QA-GNN [Yasunaga et al., 2021]

- Task: CommonsenseQA & MedQA-USMLE & OpenBookQA
- Graphs: UMLS & ConceptNet
- Idea Let's use a GNN to represent the KG structure!



QA-GNN [Yasunaga et al., 2021]



Components:

1. LM encoding: encode the question and answer set $\Rightarrow z$
2. KG retrieval: extract subgraph with entities from the QA context G
3. *QA-GNN*
   - GNN part
   - Operates on the retrieved subgraph + a node for the context z
4. *MLP*: pooled output from QA-GNN + z

But does it work?

Methods	**Test**
AristoRoBERTa	77.8
+ PG	80.2
+ QA-GNN	82.8
UnifiedQA (30x size)	87.2

Table: OpenBookQA leaderboard

Methods	**Test**
RoBERTa	72.1
+ PG	75.6
+ QA-GNN	76.1
UnifiedQA (30x size)	79.1

Table: CommonsenseQA leaderboard

Maybe a GNN is not the best way to process the extracted knowledge... (see [Huang et al., 2023])

Contents



1 Graphs

2 Graph Neural Networks

3 GNNs in NLP

4 Question Answering

5 Pretraining language models

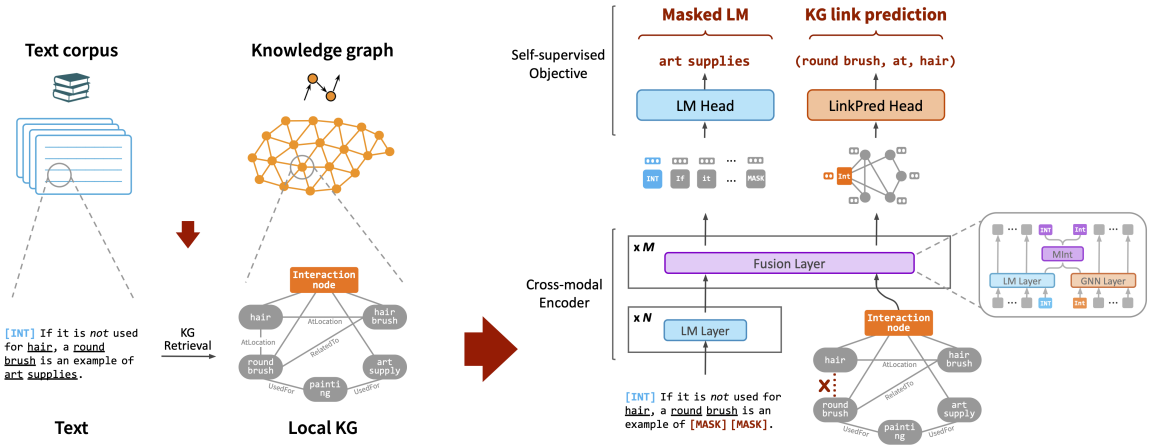
6 LLMs vs. KGs

Dragon [Yasunaga et al., 2022]

Idea: Use self-supervision to pre-train a joint language-knowledge foundation model from text and a KG.

- MLM and KG Link prediction as pre-training tasks

Dragon [Yasunaga et al., 2022]



Raw data **Pretrain DRAGON**

Joint training: $Loss = Loss_{mlm} + Loss_{link}$

Cross-modal encoder: $f((w_1, w_2 \ldots w_n), (v1, v2, \ldots, v_n))$

Contents



1 Graphs

2 Graph Neural Networks

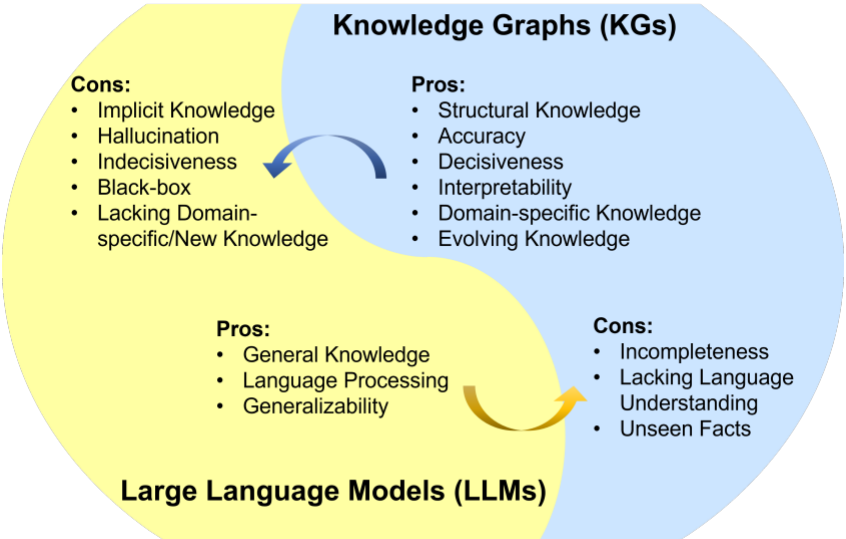
3 GNNs in NLP

4 Question Answering

5 Pretraining language models

6 LLMs vs. KGs

Large Language Models vs. Knowledge Graphs [Pan et al., 2023]



References I

Huang, Y., Li, Y., Xu, Y., Zhang, L., Gan, R., Zhang, J., and Wang, L. (2023).
MVP-tuning: Multi-view knowledge retrieval with prompt tuning for commonsense reasoning.
In Rogers, A., Boyd-Graber, J., and Okazaki, N., editors, Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 13417–13432, Toronto, Canada. Association for Computational Linguistics.

Marcheggiani, D. and Titov, I. (2017).
Encoding sentences with graph convolutional networks for semantic role labeling.
In Palmer, M., Hwa, R., and Riedel, S., editors, Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 1506–1515, Copenhagen, Denmark. Association for Computational Linguistics.

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., and Wu, X. (2023).
Unifying large language models and knowledge graphs: A roadmap.
ArXiv, abs/2306.08302.

References II

- Wu, L., Chen, Y., Shen, K., Guo, X., Gao, H., Li, S., Pei, J., and Long, B. (2021). Graph neural networks for natural language processing: A survey. CoRR, abs/2106.06090.
- Yasunaga, M., Bosselut, A., Ren, H., Zhang, X., Manning, C. D., Liang, P., and Leskovec, J. (2022). Deep bidirectional language-knowledge graph pretraining. arXiv preprint arXiv:2210.09338.
- Yasunaga, M., Ren, H., Bosselut, A., Liang, P., and Leskovec, J. (2021). QA-GNN: Reasoning with language models and knowledge graphs for question answering. In Toutanova, K., Rumshisky, A., Zettlemoyer, L., Hakkani-Tur, D., Beltagy, I., Bethard, S., Cotterell, R., Chakraborty, T., and Zhou, Y., editors, Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 535–546, Online. Association for Computational Linguistics.

References III