Graph Neural Networks for Natural Language Processing: A Survey

Other titles in Foundations and Trends® in Machine Learning

Divided Differences, Falling Factorials, and Discrete Splines: Another Look at Trend Filtering and Related Problems
Ryan J. Tibshirani
ISBN: 978-1-63828-036-1

Risk-Sensitive Reinforcement Learning via Policy Gradient Search
Prashanth L. A. and Michael C. Fu
ISBN: 978-1-63828-026-2

A Unifying Tutorial on Approximate Message Passing
Oliver Y. Feng, Ramji Venkataramanan, Cynthia Rush and Richard J. Samworth
ISBN: 978-1-63828-004-0

Learning in Repeated Auctions
Thomas Nedelec, Clément Calauzènes, Noureddine El Karoui and Vianney Perchet
ISBN: 978-1-68083-938-8

Dynamical Variational Autoencoders: A Comprehensive Review
Laurent Girin, Simon Leglaive, Xiaoyu Bie, Julien Diard, Thomas Hueber and Xavier Alameda-Pineda
ISBN: 978-1-68083-912-8

Machine Learning for Automated Theorem Proving: Learning to Solve SAT and QSATe
Sean B. Holden
ISBN: 978-1-68083-898-5

Graph Neural Networks for Natural Language Processing: A Survey

Lingfei Wu
JD.COM Silicon Valley Research Center
teddy.lfwu@gmail.com

Yu Chen
Rensselaer Polytechnic Institute
hugochan2013@gmail.com

Kai Shen
Zhejiang University
shenkai@zju.edu.cn

Xiaojie Guo
JD.COM Silicon Valley Research Center
xguo7@gmu.edu

Hanning Gao
Central China Normal University
ghnqwerty@gmail.com

Shucheng Li
Nanjing University
shuchengli@smail.nju.edu.cn

Jian Pei
Simon Fraser University
jpei@cs.sfu.ca

Bo Long
JD.COM
bo.long@jd.com

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Machine Learning

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

L. Wu *et al.*. *Graph Neural Networks for Natural Language Processing: A Survey.* Foundations and Trends® in Machine Learning, vol. 16, no. 2, pp. 119–328, 2023.

ISBN: 978-1-63828-143-6

Foundations and Trends® in Machine Learning
Volume 16, Issue 2, 2023

Editorial Scope

Topics

Foundations and Trends® in Machine Learning publishes survey and tutorial articles in the following topics:

- Adaptive control and signal processing
- Applications and case studies
- Behavioral, cognitive and neural learning
- Bayesian learning
- Classification and prediction
- Clustering
- Data mining
- Dimensionality reduction
- Evaluation
- Game theoretic learning
- Graphical models
- Independent component analysis
- Inductive logic programming
- Kernel methods
- Markov chain Monte Carlo
- Model choice
- Nonparametric methods
- Online learning
- Optimization
- Reinforcement learning
- Relational learning
- Robustness
- Spectral methods
- Statistical learning theory
- Variational inference
- Visualization

Information for Librarians

Foundations and Trends® in Machine Learning, 2023, Volume 16, 6 issues. ISSN paper version 1935-8237. ISSN online version 1935-8245. Also available as a combined paper and online subscription.

Contents

Graph Neural Networks for Natural Language Processing: A Survey

Lingfei Wu[*1], Yu Chen[*2], Kai Shen[**3], Xiaojie Guo[4], Hanning Gao[5], Shucheng Li[6], Jian Pei[7] and Bo Long[8]

[1] *JD.COM Silicon Valley Research Center, USA; teddy.lfwu@gmail.com*
[2] *Rensselaer Polytechnic Institute, USA; hugochan2013@gmail.com*
[3] *Zhejiang University, China; shenkai@zju.edu.cn*
[4] *JD.COM Silicon Valley Research Center, USA; xguo7@gmu.edu*
[5] *Central China Normal University, China; ghnqwerty@gmail.com*
[6] *National Key Lab for Novel Software Technology, Nanjing University, China; shuchengli@smail.nju.edu.cn*
[7] *Simon Fraser University, Canada; jpei@cs.sfu.ca*
[8] *JD.COM, China; bo.long@jd.com*

ABSTRACT

Deep learning has become the dominant approach in addressing various tasks in Natural Language Processing (NLP). Although text inputs are typically represented as a sequence of tokens, there is a rich variety of NLP problems that can be best expressed with a graph structure. As a result, there is a surge of interest in developing new deep learning techniques on graphs for a large number of NLP tasks. In this survey, we present a comprehensive overview on *Graph Neural Networks (GNNs) for Natural Language Processing*. We propose a new taxonomy of GNNs for NLP, which systematically organizes existing research of GNNs for NLP

[*] Both authors contributed equally to this research.
[**] This research was done when Kai Shen was an intern at JD.COM.

Lingfei Wu, Yu Chen, Kai Shen, Xiaojie Guo, Hanning Gao, Shucheng Li, Jian Pei and Bo Long (2023), “Graph Neural Networks for Natural Language Processing: A Survey”, Foundations and Trends® in Machine Learning: Vol. 16, No. 2, pp 119–328.
DOI: 10.1561/2200000096.

along three axes: graph construction, graph representation learning, and graph based encoder-decoder models. We further introduce a large number of NLP applications that exploits the power of GNNs and summarize the corresponding benchmark datasets, evaluation metrics, and open-source codes. Finally, we discuss various outstanding challenges for making the full use of GNNs for NLP as well as future research directions. To the best of our knowledge, this is the first comprehensive overview of Graph Neural Networks for Natural Language Processing.

1

Introduction

Deep learning has become the dominant approach in coping with various tasks in Natural Language Processing (NLP) today, especially when operated on large-scale text corpora. Conventionally, text sequences are considered as a bag of tokens such as BoW and TF-IDF in NLP tasks. With the recent success of Word Embeddings techniques (Mikolov *et al.*, 2013; Pennington *et al.*, 2014), sentences are typically represented as a sequence of tokens in NLP tasks. Hence, popular deep learning techniques such as recurrent neural networks (Schuster and Paliwal, 1997) and convolutional neural networks (Krizhevsky *et al.*, 2012) have been widely applied for modeling text sequence.

However, there is a rich variety of NLP problems that can be best expressed with a graph structure. For instance, the sentence structural information in text sequence (i.e. syntactic parsing trees like dependency and constituency parsing trees) can be exploited to augment original sequence data by incorporating the task-specific knowledge. Similarly, the semantic information in sequence data (i.e. semantic parsing graphs like Abstract Meaning Representation graphs and Information Extraction graphs) can be leveraged to enhance original sequence data as well. Therefore, these graph-structured data can encode complicated pairwise relationships between entity tokens for learning more informative representations.

Unfortunately, deep learning techniques that were disruptive for Euclidean data (e.g, images) or sequence data (e.g, text) are not immediately applicable to graph-structured data, due to the complexity of graph data such as irregular structure and varying size of node neighbors. As a result, this gap has driven a tide in research for deep learning on graphs, especially in development of graph neural networks (GNNs) (Wu *et al.*, 2022; Kipf and Welling, 2016; Defferrard *et al.*, 2016; Hamilton *et al.*, 2017a).

This wave of research at the intersection of deep learning on graphs and NLP has influenced a variety of NLP tasks (Liu and Wu, 2022). There has been a surge of interest in applying and developing different GNNs variants and achieved considerable success in many NLP tasks, ranging from classification tasks like sentence classification (Henaff *et al.*, 2015; Huang and Carley, 2019), semantic role labeling (Luo and Zhao, 2020; Gui *et al.*, 2019), and relation extraction (Qu *et al.*, 2020; Sahu *et al.*, 2019), to generation tasks like machine translation (Bastings *et al.*, 2017; Beck *et al.*, 2018a), question generation (Pan *et al.*, 2020; Sachan *et al.*, 2020), and summarization (Fernandes *et al.*, 2019; Yasunaga *et al.*, 2017). Despite the successes this existing research has achieved, deep learning on graphs for NLP still encounters many challenges, namely:

- Automatically transforming original text sequence data into highly graph-structured data. Such challenge is profound in NLP since most of the NLP tasks involving using the text sequences as the original inputs. Automatic graph construction from the text sequence to utilize the underlying structural information is a crucial step in utilizing graph neural networks for NLP problems.
- Properly determining graph representation learning techniques. It is critical to come up with specially-designed GNNs to learn the unique characteristics of different graph-structures data such as undirected, directed, multi-relational and heterogeneous graphs.
- Effectively modeling complex data. Such challenge is important since many NLP tasks involve learning the mapping between the

graph-based inputs and other highly structured output data such as sequences, trees, as well as graph data with multi-types in both nodes and edges.

In this survey, we will present for the first time a comprehensive overview of *Graph Neural Networks for Natural Language Processing*. Our survey is timely for both Machine Learning and NLP communities, which covers relevant and interesting topics, including automatic graph construction for NLP, graph representation learning for NLP, various advanced GNNs-based encoder-decoder models (i.e. graph2seq, graph2tree, and graph2graph) for NLP, and the applications of GNNs in various NLP tasks. We highlight our main contributions as follows:

- We propose a new taxonomy of GNNs for NLP, which systematically organizes existing research of GNNs for NLP along four axes: graph construction, graph representation learning, and graph based encoder-decoder models.
- We present the most comprehensive overview of the state-of-the-art GNNs-based approaches for various NLP tasks. We provide detailed descriptions and necessary comparisons on various graph construction approaches based on the domain knowledge and semantic space, graph representation learning approaches for various categories of graph-structures data, GNNs-based encoder-decoder models given different combinations of inputs and output data types.
- We introduce a large number of NLP applications that are exploiting the power of GNNs, including how they handle these NLP tasks along three key components (i.e., graph construction, graph representation learning, and embedding initialization), as well as providing corresponding benchmark datasets, evaluation metrics, and open-source codes.
- We outline various outstanding challenges for making the full use of GNNs for NLP and provides discussions and suggestions for fruitful and unexplored research directions.

The rest of the survey is structured as follows. Section 2 reviews the NLP problems from a graph perspective, and then briefly introduces some representative traditional graph-based methods for solving NLP problems. Section 3 elaborates basic foundations and methodologies for graph neural networks, which are a class of modern neural networks that directly operate on graph-structured data. We also provide a list of notations used throughout this survey. Section 4 focuses on introducing two major graph construction approaches, namely static graph construction and dynamic graph construction for constructing graph structured inputs in various NLP tasks. Section 5 discusses various graph representation learning techniques that are directly operated on the constructed graphs for various NLP tasks. Section 6 first introduces the typical Seq2Seq models, and then discusses two typical graph-based encoder-decoder models for NLP tasks (i.e., graph-to-tree and graph-to-graph models). Section 7 discusses 12 typical NLP applications using GNNs by providing the summary of all the applications with their sub-tasks, evaluation metrics and open-source codes. Section 8 discusses various general challenges of GNNs for NLP and pinpoints the future research directions. Finally, Section 9 summarizes the survey. The taxonomy, which systematically organizes GNN for NLP approaches along four axes: graph construction, graph representation learning, encoder-decoder models, and the applications are illustrated in Figure 1.1.

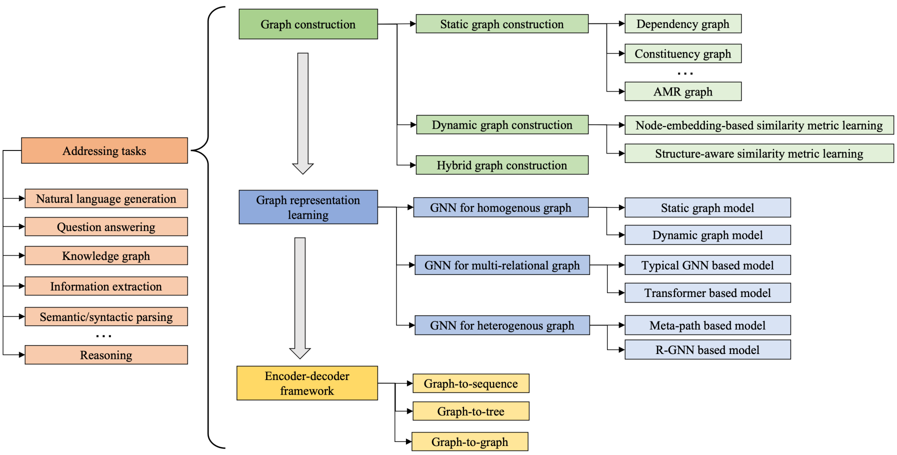

Figure 1.1: The taxonomy, which systematically organizes GNNs for NLP along four axes: graph construction, graph representation learning, encoder-decoder models, and the applications.

References

Abbe, E. (2017). "Community detection and stochastic block models: recent developments". *The Journal of Machine Learning Research*. 18(1): 6446–6531.

Allahyari, M., S. Pouriyeh, M. Assefi, S. Safaei, E. D. Trippe, J. B. Gutierrez, and K. Kochut. (2017). "Text summarization techniques: a brief survey". *arXiv preprint arXiv:1707.02268*.

Allamanis, M., M. Brockschmidt, and M. Khademi. (2018). "Learning to Represent Programs with Graphs". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=BJOFETxR-.

Alon, U., S. Brody, O. Levy, and E. Yahav. (2018). "code2seq: Generating sequences from structured representations of code". *arXiv preprint arXiv:1808.01400*.

Alvarez-Melis, D. and T. S. Jaakkola. (2016). "Tree-structured decoding with doubly-recurrent neural networks".

Amini, A., S. Gabriel, P. Lin, R. Koncel-Kedziorski, Y. Choi, and H. Hajishirzi. (2019). "Mathqa: Towards interpretable math word problem solving with operation-based formalisms". *arXiv preprint arXiv:1905.13319*.

Andor, D., C. Alberti, D. Weiss, A. Severyn, A. Presta, K. Ganchev, S. Petrov, and M. Collins. (2016). "Globally Normalized Transition-Based Neural Networks". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics. 2442–2452. DOI: 10.18653/v1/P16-1231.

Angeli, G., M. J. Johnson Premkumar, and C. D. Manning. (2015). "Leveraging Linguistic Structure For Open Domain Information Extraction". In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Beijing, China: Association for Computational Linguistics. 344–354. DOI: 10.3115/v1/P15-1034.

Atwood, J. and D. Towsley. (2016). "Diffusion-Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett. Vol. 29. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2016/file/390e982518a50e280d8e2b535462ec1f-Paper.pdf.

Ba, J. L., J. R. Kiros, and G. E. Hinton. (2016). "Layer normalization". *arXiv preprint arXiv:1607.06450*.

Bahdanau, D., K. Cho, and Y. Bengio. (2015). "Neural Machine Translation by Jointly Learning to Align and Translate". In: *3rd International Conference on Learning Representations*. Ed. by Y. Bengio and Y. LeCun.

Bai, X., Y. Chen, L. Song, and Y. Zhang. (2021). "Semantic Representation for Dialogue Modeling". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 4430–4445.

Bai, X., L. Song, and Y. Zhang. (2020). "Online Back-Parsing for AMR-to-Text Generation". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 1206–1219. DOI: 10.18653/v1/2020.emnlp-main.92.

Bansal, T., D.-C. Juan, S. Ravi, and A. McCallum. (2019). "A2N: attending to neighbors for knowledge graph inference". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 4387–4392.

Bao, J., D. Tang, N. Duan, Z. Yan, Y. Lv, M. Zhou, and T. Zhao. (2018). "Table-to-text: Describing table region with natural language". In: *Thirty-Second AAAI Conference on Artificial Intelligence*.

Barone, A. V. M. and R. Sennrich. (2017). "A parallel corpus of Python functions and documentation strings for automated code documentation and code generation". *arXiv preprint arXiv:1707.02275*.

Bastings, J., I. Titov, W. Aziz, D. Marcheggiani, and K. Sima'an. (2017). "Graph Convolutional Encoders for Syntax-aware Neural Machine Translation". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics. 1957–1967. DOI: 10.18653/v1/D17-1209.

Beck, D., G. Haffari, and T. Cohn. (2018a). "Graph-to-Sequence Learning using Gated Graph Neural Networks". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia: Association for Computational Linguistics. 273–283. DOI: 10.18653/v1/P18-1026.

Beck, D., G. Haffari, and T. Cohn. (2018b). "Graph-to-Sequence Learning using Gated Graph Neural Networks". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 273–283.

Bengio, S., O. Vinyals, N. Jaitly, and N. Shazeer. (2015). "Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks". In: *Advances in Neural Information Processing Systems 28: Annual Conference on Neural Information Processing Systems 2015, December 7-12, 2015, Montreal, Quebec, Canada*. Ed. by C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett. 1171–1179. URL: https://proceedings.neurips.cc/paper/2015/hash/e995f98d56967d946471af29d7bf99f1-Abstract.html.

Blei, D. M., T. L. Griffiths, and M. I. Jordan. (2010). "The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies". *Journal of the ACM (JACM)*. 57(2): 1–30.

Blei, D. M., A. Y. Ng, and M. I. Jordan. (2003). "Latent dirichlet allocation". *the Journal of machine Learning research*. 3: 993–1022.

Blitzer, J., M. Dredze, and F. Pereira. (2007). "Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification". In: *Proceedings of the 45th annual meeting of the association of computational linguistics*. 440–447.

Bogin, B., J. Berant, and M. Gardner. (2019a). "Representing Schema Structure with Graph Neural Networks for Text-to-SQL Parsing". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 4560–4565. DOI: 10.18653/v1/P19-1448.

Bogin, B., M. Gardner, and J. Berant. (2019b). "Global Reasoning over Database Structures for Text-to-SQL Parsing". 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)*. Hong Kong, China: Association for Computational Linguistics. 3659–3664. DOI: 10.18653/v1/D19-1378.

Bordes, A., N. Usunier, S. Chopra, and J. Weston. (2015). "Large-scale simple question answering with memory networks". *arXiv preprint arXiv:1506.02075*.

Bordes, A., N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. (2013). "Translating embeddings for modeling multi-relational data". In: *Neural Information Processing Systems (NIPS)*. 1–9.

Bowman, S. R., G. Angeli, C. Potts, and C. D. Manning. (2015). "A large annotated corpus for learning natural language inference". *arXiv preprint arXiv:1508.05326*.

Brown, T. B., B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, *et al.* (2020). "Language models are few-shot learners". *arXiv preprint arXiv:2005.14165*.

Budzianowski, P., T.-H. Wen, B.-H. Tseng, I. Casanueva, S. Ultes, O. Ramadan, and M. Gasic. (2018). "MultiWOZ-A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling". In: *EMNLP*.

Cai, D. and W. Lam. (2020a). "AMR Parsing via Graph-Sequence Iterative Inference". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 1290–1301. DOI: 10.18653/v1/2020.acl-main.119.

Cai, D. and W. Lam. (2020b). "Graph Transformer for Graph-to-Sequence Learning". In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence*. AAAI Press. 7464–7471.

Cai, D. and W. Lam. (2020c). "Graph Transformer for Graph-to-Sequence Learning". *Proceedings of the AAAI Conference on Artificial Intelligence*. 34(05): 7464–7471. DOI: 10.1609/aaai.v34i05.6243.

Cai, S. and K. Knight. (2013). "Smatch: an Evaluation Metric for Semantic Feature Structures". In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Sofia, Bulgaria: Association for Computational Linguistics. 748–752. URL: https://www.aclweb.org/anthology/P13-2131.

Campos, J. A., A. Otegi, A. Soroa, J. Deriu, M. Cieliebak, and E. Agirre. (2019). "Conversational qa for faqs". In: *3rd Conversational AI:"Today's Practice and Tomorrow's Potential" workshop*.

Cao, N. D., W. Aziz, and I. Titov. (2019a). "Question Answering by Reasoning Across Documents with Graph Convolutional Networks". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*. Ed. by J. Burstein, C. Doran, and T. Solorio. Association for Computational Linguistics. 2306–2317.

Cao, Y., Z. Liu, C. Li, J. Li, and T.-S. Chua. (2019b). "Multi-Channel Graph Neural Network for Entity Alignment". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 1452–1461.

Cao, Y., M. Fang, and D. Tao. (2019c). “BAG: Bi-directional Attention Entity Graph Convolutional Network for Multi-hop Reasoning Question Answering”. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*. Association for Computational Linguistics. 357–362.

Casanueva, I., P. Budzianowski, P.-H. Su, N. Mrkšic, T.-H. Wen, S. Ultes, L. Rojas-Barahona, S. Young, and M. Gašic. (2017). “A Benchmarking Environment for Reinforcement Learning Based Task Oriented Dialogue Management”. *stat*. 1050: 29.

Chen, C., Z. Teng, and Y. Zhang. (2020a). “Inducing target-specific latent structures for aspect sentiment classification”. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 5596–5607.

Chen, D., A. Fisch, J. Weston, and A. Bordes. (2017a). “Reading wikipedia to answer open-domain questions”. *arXiv preprint arXiv:1704.00051*.

Chen, H., X. Liu, D. Yin, and J. Tang. (2017b). “A survey on dialogue systems: Recent advances and new frontiers”. *Acm Sigkdd Explorations Newsletter*. 19(2): 25–35.

Chen, J., Q. Chen, X. Liu, H. Yang, D. Lu, and B. Tang. (2018a). “The bq corpus: A large-scale domain-specific chinese corpus for sentence semantic equivalence identification”. In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 4946–4951.

Chen, L., B. Lv, C. Wang, S. Zhu, B. Tan, and K. Yu. (2020b). “Schema-Guided Multi-Domain Dialogue State Tracking with Graph Attention Neural Networks”. In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*. AAAI Press. 7521–7528.

Chen, L., B. Tan, S. Long, and K. Yu. (2018b). “Structured Dialogue Policy with Graph Neural Networks”. In: *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018*. Ed. by E. M. Bender, L. Derczynski, and P. Isabelle. Association for Computational Linguistics. 1257–1268.

Chen, L., Y. Zhao, B. Lyu, L. Jin, Z. Chen, S. Zhu, and K. Yu. (2020c). “Neural Graph Matching Networks for Chinese Short Text Matching”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*. Ed. by D. Jurafsky, J. Chai, N. Schluter, and J. R. Tetreault. Association for Computational Linguistics. 6152–6158.

Chen, Q., X. Zhu, Z.-H. Ling, D. Inkpen, and S. Wei. (2017c). “Neural natural language inference models enhanced with external knowledge”. *arXiv preprint arXiv:1711.04289*.

Chen, W., Y. Su, X. Yan, and W. Y. Wang. (2020d). “KGPT: Knowledge-Grounded Pre-Training for Data-to-Text Generation”. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*. Association for Computational Linguistics. 8635–8648.

Chen, X., C. Sun, J. Wang, S. Li, L. Si, M. Zhang, and G. Zhou. (2020e). “Aspect Sentiment Classification with Document-level Sentiment Preference Modeling”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 3667–3677. DOI: 10.18653/v1/2020.acl-main.338.

Chen, Y., L. Wu, and M. J. Zaki. (2019). “Bidirectional Attentive Memory Networks for Question Answering over Knowledge Bases”. In: *NAACL-HLT (1)*.

Chen, Y., L. Wu, and M. J. Zaki. (2020f). “Iterative Deep Graph Learning for Graph Neural Networks: Better and Robust Node Embeddings”. In: *Proceedings of the 34th Conference on Neural Information Processing Systems*.

Chen, Y., L. Wu, and M. J. Zaki. (2020g). “Toward Subgraph Guided Knowledge Graph Question Generation with Graph Neural Networks”. *arXiv preprint arXiv:2004.06015*.

Chen, Y., L. Wu, and M. J. Zaki. (2020h). "GraphFlow: Exploiting Conversation Flow with Graph Neural Networks for Conversational Machine Comprehension". In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*. ijcai.org. 1230–1236.

Chen, Y., L. Wu, and M. J. Zaki. (2020i). "Reinforcement Learning Based Graph-to-Sequence Model for Natural Question Generation". In: *Proceedings of the 8th International Conference on Learning Representations*.

Chen, Y. and M. J. Zaki. (2017). "Kate: K-competitive autoencoder for text". In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 85–94.

Cho, K., B. van Merrienboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. (2014). "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*. Ed. by A. Moschitti, B. Pang, and W. Daelemans. ACL. 1724–1734. DOI: 10.3115/v1/d14-1179.

Choi, E., H. He, M. Iyyer, M. Yatskar, W. T. Yih, Y. Choi, P. Liang, and L. Zettlemoyer. (2020). "QUAC: Question answering in context". In: *2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*. Association for Computational Linguistics. 2174–2184.

Choromanski, K., V. Likhosherstov, D. Dohan, X. Song, A. Gane, T. Sarlós, P. Hawkins, J. Davis, A. Mohiuddin, L. Kaiser, D. Belanger, L. Colwell, and A. Weller. (2021). "Rethinking Attention with Performers". In: *International Conference on Learning Representations*.

Christensen, J., S. Soderland, O. Etzioni, *et al.* (2013). "Towards coherent multi-document summarization". In: *Proceedings of the 2013 conference of the North American chapter of the association for computational linguistics: Human language technologies*. 1163–1173.

Christopoulou, F., M. Miwa, and S. Ananiadou. (2019). "Connecting the Dots: Document-level Neural Relation Extraction with Edge-oriented Graphs". 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)*. Hong Kong, China: Association for Computational Linguistics. 4925–4936. DOI: 10.18653/v1/D19-1498.

Collins-Thompson, K. and J. Callan. (2005). "Query expansion using random walk models". In: *Proceedings of the 14th ACM international conference on Information and knowledge management*. 704–711.

Cui, L., Y. Wu, S. Liu, Y. Zhang, and M. Zhou. (2020a). "MuTual: A Dataset for Multi-Turn Dialogue Reasoning". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 1406–1416.

Cui, P., L. Hu, and Y. Liu. (2020b). "Enhancing Extractive Text Summarization with Topic-Aware Graph Neural Networks". *arXiv preprint arXiv:2010. 06253*.

Cui, S., B. Yu, T. Liu, Z. Zhang, X. Wang, and J. Shi. (2020c). "Edge-Enhanced Graph Convolution Networks for Event Detection with Syntactic Relation". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 2329–2339. DOI: 10.18653/v1/2020.findings-emnlp.211.

Cui, Y., Z. Chen, S. Wei, S. Wang, T. Liu, and G. Hu. (2017). "Attention-over-Attention Neural Networks for Reading Comprehension". In: *ACL (1)*.

Dahl, D. A., M. Bates, M. Brown, W. Fisher, K. Hunicke-Smith, D. Pallett, C. Pao, A. Rudnicky, and E. Shriberg. (1994). "Expanding the Scope of the ATIS Task: The ATIS-3 Corpus". In: *Human Language Technology: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994*. URL: https://www.aclweb.org/anthology/H94-1010.

Damonte, M. and S. B. Cohen. (2019). "Structural Neural Encoders for AMR-to-text Generation". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics. 3649–3658. DOI: 10.18653/v1/N19-1366.

De Cao, N., W. Aziz, and I. Titov. (2018). "Question answering by reasoning across documents with graph convolutional networks".

Defferrard, M., X. Bresson, and P. Vandergheynst. (2016). "Convolutional neural networks on graphs with fast localized spectral filtering". *Advances in neural information processing systems*. 29.

Dettmers, T., P. Minervini, P. Stenetorp, and S. Riedel. (2018a). "Convolutional 2d knowledge graph embeddings". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1.

Dettmers, T., M. Pasquale, S. Pontus, and S. Riedel. (2018b). "Convolutional 2D Knowledge Graph Embeddings". In: *Proceedings of the 32th AAAI Conference on Artificial Intelligence*. 1811–1818. URL: https://arxiv.org/abs/1707.01476.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. (2018). "Bert: Pre-training of deep bidirectional transformers for language understanding". *arXiv preprint arXiv:1810.04805*.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics. 4171–4186. DOI: 10.18653/v1/N19-1423.

Ding, M., C. Zhou, Q. Chen, H. Yang, and J. Tang. (2019a). "Cognitive Graph for Multi-Hop Reading Comprehension at Scale". In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*. Ed. by A. Korhonen, D. R. Traum, and L. Màrquez. Association for Computational Linguistics. 2694–2703.

Ding, R., P. Xie, X. Zhang, W. Lu, L. Li, and L. Si. (2019b). "A Neural Multi-digraph Model for Chinese NER with Gazetteers". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 1462–1467. DOI: 10.18653/v1/P19-1141.

Do, B.-N. and I. Rehbein. (2020). "Neural Reranking for Dependency Parsing: An Evaluation". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 4123–4133. DOI: 10.18653/v1/2020.acl-main.379.

Do, K., T. Tran, and S. Venkatesh. (2019). "Graph transformation policy network for chemical reaction prediction". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 750–760.

Dong, L. and M. Lapata. (2016). "Language to Logical Form with Neural Attention". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics. 33–43. DOI: 10.18653/v1/P16-1004.

Dong, L., F. Wei, C. Tan, D. Tang, M. Zhou, and K. Xu. (2014). "Adaptive recursive neural network for target-dependent twitter sentiment classification". In: *Proceedings of the 52nd annual meeting of the association for computational linguistics (volume 2: Short papers)*. 49–54.

Dozat, T. and C. D. Manning. (2016). "Deep biaffine attention for neural dependency parsing". *arXiv preprint arXiv:1611.01734*.

Du, X., J. Shao, and C. Cardie. (2017). "Learning to Ask: Neural Question Generation for Reading Comprehension". In: *ACL (1)*.

Dua, D., Y. Wang, P. Dasigi, G. Stanovsky, S. Singh, and M. Gardner. (2019). "DROP: A Reading Comprehension Benchmark Requiring Discrete Reasoning Over Paragraphs". In: *Proceedings of NAACL-HLT*. 2368–2378.

Dyer, C., A. Kuncoro, M. Ballesteros, and N. A. Smith. (2016). "Recurrent Neural Network Grammars". In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. San Diego, California: Association for Computational Linguistics. 199–209. DOI: 10.18653/v1/N16-1024.

Edouard, A., E. Cabrio, S. Tonelli, and N. Le-Thanh. (2017). "Graph-based Event Extraction from Twitter". In: *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*. Varna, Bulgaria: INCOMA Ltd. 222–230. DOI: 10.26615/978-954-452-049-6_031.

Elliott, D., S. Frank, K. Sima'an, and L. Specia. (2016). "Multi30K: Multilingual English-German Image Descriptions". In: *Proceedings of the 5th Workshop on Vision and Language*. Berlin, Germany: Association for Computational Linguistics. 70–74. DOI: 10.18653/v1/W16-3210.

Eric, M., R. Goel, S. Paul, A. Sethi, S. Agarwal, S. Gao, A. Kumar, A. K. Goyal, P. Ku, and D. Hakkani-Tür. (2020). "MultiWOZ 2.1: A Consolidated Multi-Domain Dialogue Dataset with State Corrections and State Tracking Baselines". In: *LREC*.

Eriguchi, A., K. Hashimoto, and Y. Tsuruoka. (2016). "Tree-to-Sequence Attentional Neural Machine Translation". In: *ACL (1)*.

Erkan, G. (2006). "Language model-based document clustering using random walks". In: *Proceedings of the Human Language Technology Conference of the NAACL, Main Conference*. 479–486.

Fabbri, A. R., I. Li, T. She, S. Li, and D. R. Radev. (2019). "Multi-news: A large-scale multi-document summarization dataset and abstractive hierarchical model". *arXiv preprint arXiv:1906.01749*.

Fan, S., J. Zhu, X. Han, C. Shi, L. Hu, B. Ma, and Y. Li. (2019). "Metapath-guided heterogeneous graph neural network for intent recommendation". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2478–2486.

Fancellu, F., S. Gilroy, A. Lopez, and M. Lapata. (2019). "Semantic graph parsing with recurrent neural network DAG grammars". 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)*. 2769–2778.

Fang, Y., S. Sun, Z. Gan, R. Pillai, S. Wang, and J. Liu. (2020a). "Hierarchical Graph Network for Multi-hop Question Answering". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 8823–8838.

Fang, Y., S. Sun, Z. Gan, R. Pillai, S. Wang, and J. Liu. (2020b). "Hierarchical Graph Network for Multi-hop Question Answering". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*. Ed. by B. Webber, T. Cohn, Y. He, and Y. Liu. Association for Computational Linguistics. 8823–8838.

Fei, H., M. Zhang, F. Li, and D. Ji. (2020). "Cross-lingual semantic role labeling with model transfer". *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 28: 2427–2437.

Feng, Y., X. Chen, B. Y. Lin, P. Wang, J. Yan, and X. Ren. (2020a). "Scalable Multi-Hop Relational Reasoning for Knowledge-Aware Question Answering". *arXiv preprint arXiv:2005.00646*.

Feng, Y., X. Chen, B. Y. Lin, P. Wang, J. Yan, and X. Ren. (2020b). "Scalable Multi-Hop Relational Reasoning for Knowledge-Aware Question Answering". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing, EMNLP 2020, Online, November 16-20, 2020*. Ed. by B. Webber, T. Cohn, Y. He, and Y. Liu. Association for Computational Linguistics. 1295–1309.

Fernandes, P., M. Allamanis, and M. Brockschmidt. (2019). "Structured Neural Summarization". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=H1ersoRqtm.

Ferreira, D. and A. Freitas. (2020). "Premise Selection in Natural Language Mathematical Texts". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 7365–7374. DOI: 10.18653/v1/2020.acl-main.657.

Fey, M. and J. E. Lenssen. (2019). "Fast graph representation learning with PyTorch Geometric". *arXiv preprint arXiv:1903.02428*.

Flanigan, J., S. Thomson, J. G. Carbonell, C. Dyer, and N. A. Smith. (2014). "A discriminative graph-based parser for the abstract meaning representation". In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 1426–1436.

Franceschi, L., M. Niepert, M. Pontil, and X. He. (2019). "Learning Discrete Structures for Graph Neural Networks". In: *Proceedings of the 36th International Conference on Machine*. Vol. 97. 1972–1982.

Fu, Q., L. Song, W. Du, and Y. Zhang. (2021). "End-to-End AMR Coreference Resolution". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 4204–4214.

Fu, T.-J., P.-H. Li, and W.-Y. Ma. (2019). "GraphRel: Modeling text as relational graphs for joint entity and relation extraction". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 1409–1418.

Gao, H., L. Wu, P. Hu, and F. Xu. (2020). "RDF-to-Text Generation with Graph-augmented Structural Neural Encoders". In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. International Joint Conferences on Artificial Intelligence Organization. 3030–3036.

Gao, H., Y. Chen, and S. Ji. (2019). "Learning graph pooling and hybrid convolutional operations for text representations". In: *The World Wide Web Conference*. 2743–2749.

Gao, Q. and S. Vogel. (2011). "Corpus expansion for statistical machine translation with semantic role label substitution rules". In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. 294–298.

Gardent, C., A. Shimorina, S. Narayan, and L. Perez-Beltrachini. (2017). "Creating Training Corpora for NLG Micro-Planners". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada: Association for Computational Linguistics. 179–188. DOI: 10.18653/v1/P17-1017.

Gehring, J., M. Auli, D. Grangier, D. Yarats, and Y. N. Dauphin. (2017). "Convolutional sequence to sequence learning". In: *International Conference on Machine Learning*. PMLR. 1243–1252.

Gehrmann, S., Y. Deng, and A. M. Rush. (2018). "Bottom-up abstractive summarization". *arXiv preprint arXiv:1808.10792*.

Ghosal, D., D. Hazarika, A. Roy, N. Majumder, R. Mihalcea, and S. Poria. (2020). "KinGDOM: Knowledge-Guided DOMain Adaptation for Sentiment Analysis". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 3198–3210. DOI: 10.18653/v1/2020.acl-main.292.

Gilmer, J., S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl. (2017). "Neural message passing for quantum chemistry". In: *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org. 1263–1272.

Goldberg, A. B. and X. Zhu. (2006). "Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization". In: *Proceedings of TextGraphs: The first workshop on graph based methods for natural language processing*. 45–52.

Gómez-Bombarelli, R., J. N. Wei, D. Duvenaud, J. M. Hernández-Lobato, B. Sánchez-Lengeling, D. Sheberla, J. Aguilera-Iparraguirre, T. D. Hirzel, R. P. Adams, and A. Aspuru-Guzik. (2018). "Automatic chemical design using a data-driven continuous representation of molecules". *ACS central science*. 4(2): 268–276.

Gu, J., Z. Lu, H. Li, and V. O. Li. (2016). "Incorporating copying mechanism in sequence-to-sequence learning". *arXiv preprint arXiv:1603.06393*.

Gui, T., Y. Zou, Q. Zhang, M. Peng, J. Fu, Z. Wei, and X. Huang. (2019). "A Lexicon-Based Graph Neural Network for Chinese NER". 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)*. Hong Kong, China: Association for Computational Linguistics. 1040–1050. DOI: 10.18653/v1/D19-1096.

Guo, X., L. Wu, and L. Zhao. (2018). "Deep graph translation". *arXiv preprint arXiv:1805.09980*.

Guo, X., L. Zhao, C. Nowzari, S. Rafatirad, H. Homayoun, and S. M. P. Dinakarrao. (2019a). "Deep multi-attributed graph translation with node-Edge Co-evolution". In: *2019 IEEE International Conference on Data Mining (ICDM)*. IEEE. 250–259.

Guo, Z., Y. Zhang, and W. Lu. (2019b). "Attention Guided Graph Convolutional Networks for Relation Extraction". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 241–251. DOI: 10.18653/v1/P19-1024.

Guo, Z., Y. Zhang, Z. Teng, and W. Lu. (2019c). "Densely Connected Graph Convolutional Networks for Graph-to-Sequence Learning". *Transactions of the Association for Computational Linguistics*. 7(Mar.): 297–312. DOI: 10.1162/tacl_a_00269.

Gupta, S., S. Kenkre, and P. Talukdar. (2019). "CaRe: Open Knowledge Graph Embeddings". 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)*. Hong Kong, China: Association for Computational Linguistics. 378–388. DOI: 10.18653/v1/D19-1036.

Haghighi, A., A. Y. Ng, and C. D. Manning. (2005). "Robust textual inference via graph matching". In: *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*. 387–394.

Hamilton, W., Z. Ying, and J. Leskovec. (2017a). "Inductive representation learning on large graphs". In: *Advances in Neural Information Processing Systems*. 1024–1034.

Hamilton, W. L., R. Ying, and J. Leskovec. (2017b). "Representation learning on graphs: Methods and applications". *arXiv preprint arXiv:1709.05584*.

Han, J., B. Cheng, and X. Wang. (2020). "Open Domain Question Answering based on Text Enhanced Knowledge Graph with Hyperedge Infusion". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 1475–1481. DOI: 10.18653/v1/2020.findings-emnlp.133.

Hashimoto, K. and Y. Tsuruoka. (2017). "Neural Machine Translation with Source-Side Latent Graph Parsing". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Copenhagen, Denmark: Association for Computational Linguistics. 125–135. DOI: 10.18653/v1/D17-1012.

Haveliwala, T. H. (2002). "Topic-Sensitive PageRank". In: *Proceedings of the 11th International Conference on World Wide Web. WWW '02*. Honolulu, Hawaii, USA: Association for Computing Machinery. 517–526. DOI: 10.1145/511446.511513.

He, B., D. Zhou, J. Xiao, X. Jiang, Q. Liu, N. J. Yuan, and T. Xu. (2020). "Integrating Graph Contextualized Knowledge into Pre-trained Language Models". In: *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020*. Vol. EMNLP 2020. *Findings of ACL*. Association for Computational Linguistics. 2281–2290.

He, L., K. Lee, O. Levy, and L. Zettlemoyer. (2018). "Jointly predicting predicates and arguments in neural semantic role labeling". *arXiv preprint arXiv:1805.04787*.

He, L., K. Lee, M. Lewis, and L. Zettlemoyer. (2017). "Deep semantic role labeling: What works and what's next". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 473–483.

Henaff, M., J. Bruna, and Y. LeCun. (2015). "Deep convolutional networks on graph-structured data". *arXiv preprint arXiv:1506.05163*.

Hermann, K. M., T. Kociskỳ, E. Grefenstette, L. Espeholt, W. Kay, M. Suleyman, and P. Blunsom. (2015). "Teaching Machines to Read and Comprehend". In: *NIPS*.

Hersh, W., C. Buckley, T. Leone, and D. Hickam. (1994). "OHSUMED: An interactive retrieval evaluation and new large test collection for research". In: *SIGIR'94*. Springer. 192–201.

Hochreiter, S. and J. Schmidhuber. (1997). "Long short-term memory". *Neural computation*. 9(8): 1735–1780.

Hoffman, M. D., D. M. Blei, C. Wang, and J. Paisley. (2013). "Stochastic variational inference". *The Journal of Machine Learning Research*. 14(1): 1303–1347.

Hofmann, T. (1999). "Probabilistic latent semantic indexing". In: *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*. 50–57.

Hu, B., Z. Lu, H. Li, and Q. Chen. (2014). "Convolutional Neural Network Architectures for Matching Natural Language Sentences". In: *NIPS*.

Hu, J., Q. Fang, S. Qian, and C. Xu. (2020a). "Multi-Modal Attentive Graph Pooling Model for Community Question Answer Matching". In: *Proceedings of the 28th ACM International Conference on Multimedia. MM '20*. Seattle, WA, USA: Association for Computing Machinery. 3505–3513. DOI: 10.1145/3394171.3413711.

Hu, J., Q. Fang, S. Qian, and C. Xu. (2020b). "Multi-modal Attentive Graph Pooling Model for Community Question Answer Matching". In: *Proceedings of the 28th ACM International Conference on Multimedia*. 3505–3513.

Hu, J., S. Qian, Q. Fang, and C. Xu. (2018). "Attentive Interactive Convolutional Matching for Community Question Answering in Social Multimedia". In: *Proceedings of the 26th ACM International Conference on Multimedia. MM '18*. Seoul, Republic of Korea: Association for Computing Machinery. 456–464. DOI: 10.1145/3240508.3240626.

Hu, J., S. Qian, Q. Fang, and C. Xu. (2019a). "Hierarchical Graph Semantic Pooling Network for Multi-Modal Community Question Answer Matching". In: *Proceedings of the 27th ACM International Conference on Multimedia. MM '19*. Nice, France: Association for Computing Machinery. 1157–1165. DOI: 10.1145/3343031.3350966.

Hu, J., S. Qian, Q. Fang, and C. Xu. (2019b). "Hierarchical graph semantic pooling network for multi-modal community question answer matching". In: *Proceedings of the 27th ACM International Conference on Multimedia*. 1157–1165.

Hu, L., T. Yang, C. Shi, H. Ji, and X. Li. (2019c). "Heterogeneous Graph Attention Networks for Semi-supervised Short Text Classification". 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 2019, Hong Kong, China, November 3-7, 2019*. Ed. by K. Inui, J. Jiang, V. Ng, and X. Wan. Association for Computational Linguistics. 4820–4829.

Hu, W., B. Liu, J. Gomes, M. Zitnik, P. Liang, V. Pande, and J. Leskovec. (2019d). "Strategies for pre-training graph neural networks". *arXiv preprint arXiv:1905.12265*.

Hu, W., Z. Chan, B. Liu, D. Zhao, J. Ma, and R. Yan. (2019e). "GSN: A Graph-Structured Network for Multi-Party Dialogues". In: *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence, IJCAI 2019, Macao, China, August 10-16, 2019*. Ed. by S. Kraus. ijcai.org. 5010–5016.

Hu, Z., Y. Dong, K. Wang, and Y. Sun. (2020c). "Heterogeneous graph transformer". In: *Proceedings of The Web Conference 2020*. 2704–2710.

Huang, B. and K. Carley. (2018). "Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 1091–1096. DOI: 10.18653/v1/D18-1136.

Huang, B. and K. Carley. (2019). "Syntax-Aware Aspect Level Sentiment Classification with Graph Attention Networks". 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)*. Hong Kong, China: Association for Computational Linguistics. 5469–5477. DOI: 10.18653/v1/D19-1549. (Accessed on 12/26/2020).

Huang, D., P. Chen, R. Zeng, Q. Du, M. Tan, and C. Gan. (2020a). "Location-Aware Graph Convolutional Networks for Video Question Answering". In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence*. AAAI Press. 11021–11028.

Huang, L., D. Ma, S. Li, X. Zhang, and H. Wang. (2019). "Text Level Graph Neural Network for Text Classification". 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 2019, Hong Kong, China, November 3-7, 2019*. Ed. by K. Inui, J. Jiang, V. Ng, and X. Wan. Association for Computational Linguistics. 3442–3448.

Huang, L., L. Wu, and L. Wang. (2020b). "Knowledge Graph-Augmented Abstractive Summarization with Semantic-Driven Cloze Reward". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 5094–5107. DOI: 10.18653/v1/2020.acl-main.457.

Hughes, T. and D. Ramage. (2007). "Lexical semantic relatedness with random graph walks". In: *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*. 581–589.

Huo, S., T. Ma, J. Chen, M. Chang, L. Wu, and M. Witbrock. (2019). "Graph Enhanced Cross-Domain Text-to-SQL Generation". In: *Proceedings of the Thirteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-13)*. Hong Kong: Association for Computational Linguistics. 159–163. DOI: 10.18653/v1/D19-5319.

Isonuma, M., J. Mori, D. Bollegala, and I. Sakata. (2020). "Tree-Structured Neural Topic Model". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 800–806. DOI: 10.18653/v1/2020.acl-main.73.

Iyer, S., I. Konstas, A. Cheung, and L. Zettlemoyer. (2016). "Summarizing source code using a neural attention model". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 2073–2083.

Ji, S., S. Pan, E. Cambria, P. Marttinen, and P. S. Yu. (2020). "A survey on knowledge graphs: Representation, acquisition and applications". *arXiv preprint arXiv:2002.00388*.

Ji, T., Y. Wu, and M. Lan. (2019). "Graph-based Dependency Parsing with Graph Neural Networks". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 2475–2485. DOI: 10.18653/v1/P19-1237.

Jia, R. and P. Liang. (2016). "Data Recombination for Neural Semantic Parsing". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Berlin, Germany: Association for Computational Linguistics. 12–22. DOI: 10.18653/v1/P16-1002.

Jia, R., Y. Cao, H. Tang, F. Fang, C. Cao, and S. Wang. (2020). "Neural Extractive Summarization with Hierarchical Attentive Heterogeneous Graph Network". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 3622–3631. DOI: 10.18653/v1/2020.emnlp-main.295.

Jiang, L., M. Yu, M. Zhou, X. Liu, and T. Zhao. (2011). "Target-dependent twitter sentiment classification". In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies*. 151–160.

Jiang, Q., L. Chen, R. Xu, X. Ao, and M. Yang. (2019). "A challenge dataset and effective models for aspect-based sentiment analysis". 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)*. 6281–6286.

Jin, H., L. Hou, J. Li, and T. Dong. (2019). "Fine-Grained Entity Typing via Hierarchical Multi Graph Convolutional Networks". 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)*. Hong Kong, China: Association for Computational Linguistics. 4969–4978. DOI: 10.18653/v1/D19-1502.

Jin, H., T. Wang, and X. Wan. (2020a). "SemSUM: Semantic Dependency Guided Neural Abstractive Summarization". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 05. 8026–8033.

Jin, H., T. Wang, and X. Wan. (2020b). "SemSUM: Semantic Dependency Guided Neural Abstractive Summarization". *Proceedings of the AAAI Conference on Artificial Intelligence*. 34(05): 8026–8033. DOI: 10.1609/aaai.v34i05.6312.

Jin, L. and D. Gildea. (2020). "Generalized Shortest-Paths Encoders for AMR-to-Text Generation". In: *Proceedings of the 28th International Conference on Computational Linguistics*. Barcelona, Spain (Online): International Committee on Computational Linguistics. 2004–2013. URL: https://www.aclweb.org/anthology/2020.coling-main.181.

Joulin, A., É. Grave, P. Bojanowski, and T. Mikolov. (2017). "Bag of Tricks for Efficient Text Classification". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. 427–431.

Kalofolias, V. and N. Perraudin. (2019). "Large Scale Graph Learning From Smooth Signals". In: *7th International Conference on Learning Representations*.

Kapanipathi, P., V. Thost, S. Patel, S. Whitehead, I. Abdelaziz, A. Balakrishnan, M. Chang, K. P. Fadnis, R. C. Gunasekara, B. Makni, N. Mattei, K. Talamadupula, and A. Fokoue. (2020). "Infusing Knowledge into the Textual Entailment Task Using Graph Convolutional Networks". In: *AAAI*.

Katharopoulos, A., A. Vyas, N. Pappas, and F. Fleuret. (2020). "Transformers are rnns: Fast autoregressive transformers with linear attention". In: *International Conference on Machine Learning*. PMLR. 5156–5165.

Khot, T., A. Sabharwal, and P. Clark. (2018). "Scitail: A textual entailment dataset from science question answering". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1.

Kim, Y. (2014). “Convolutional Neural Networks for Sentence Classification”. In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL*. Ed. by A. Moschitti, B. Pang, and W. Daelemans. ACL. 1746–1751. DOI: 10.3115/v1/d14-1181.

Kingsbury, P. R. and M. Palmer. (2002). “From TreeBank to PropBank.” In: *LREC*. Citeseer. 1989–1993.

Kiperwasser, E. and Y. Goldberg. (2016). “Simple and Accurate Dependency Parsing Using Bidirectional LSTM Feature Representations”. *Transactions of the Association for Computational Linguistics*. 4: 313–327. DOI: 10.1162/tacl_a_00101.

Kipf, T. N. and M. Welling. (2016). “Semi-supervised classification with graph convolutional networks”. *arXiv preprint arXiv:1609.02907*.

Kiros, R., Y. Zhu, R. Salakhutdinov, R. S. Zemel, R. Urtasun, A. Torralba, and S. Fidler. (2015). “Skip-Thought Vectors”. In: *NIPS*.

Klicpera, J., A. Bojchevski, and S. Günnemann. (2019). “Combining Neural Networks with Personalized PageRank for Classification on Graphs”. In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=H1gL-2A9Ym.

Knight, K., B. Badarau, L. Banarescu, C. Bonial, M. Bardocz, K. Griffitt, U. Hermjakob, D. Marcu, M. Palmer, T. O’Gorman, *et al.* (2017). “Abstract Meaning Representation (AMR) Annotation Release 2.0”. *Tech. rep.* Technical Report LDC2017T10, Linguistic Data Consortium, Philadelphia, PA, June.

Knight, K., L. Baranescu, C. Bonial, M. Georgescu, K. Griffitt, U. Hermjakob, D. Marcu, M. Palmer, and N. Schneider. (2014). “Abstract meaning representation (AMR) annotation release 1.0 LDC2014T12”. *Web Download. Philadelphia: Linguistic Data Consortium*.

Koncel-Kedziorski, R., S. Roy, A. Amini, N. Kushman, and H. Hajishirzi. (2016). “MAWPS: A math word problem repository”. In: *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1152–1157.

Koncel-Kedziorski, R., D. Bekal, Y. Luan, M. Lapata, and H. Hajishirzi. (2019). "Text Generation from Knowledge Graphs with Graph Transformers". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics. 2284–2293. DOI: 10.18653/v1/N19-1238.

Krizhevsky, A., I. Sutskever, and G. E. Hinton. (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Ed. by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger. Vol. 25. Curran Associates, Inc. URL: https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Kumar, V., Y. Hua, G. Ramakrishnan, G. Qi, L. Gao, and Y.-F. Li. (2019). "Difficulty-controllable multi-hop question generation from knowledge graphs". In: *International Semantic Web Conference*. Springer. 382–398.

Lafferty, J. D., A. McCallum, and F. C. N. Pereira. (2001). "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data". In: *Proceedings of the Eighteenth International Conference on Machine Learning (ICML 2001), Williams College, Williamstown, MA, USA, June 28 - July 1, 2001*. Ed. by C. E. Brodley and A. P. Danyluk. Morgan Kaufmann. 282–289.

Lan, Z., M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut. (2019). "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations". In: *International Conference on Learning Representations*.

Lang, K. (1995). "Newsweeder: Learning to filter netnews". In: *Machine Learning Proceedings 1995*. Elsevier. 331–339.

Larochelle, H. and S. Lauly. (2012). "A neural autoregressive topic model". *Advances in Neural Information Processing Systems*. 25: 2708–2716.

Le, Q. and T. Mikolov. (2014). "Distributed representations of sentences and documents". In: *International conference on machine learning*. PMLR. 1188–1196.

LeClair, A., S. Haque, L. Wu, and C. McMillan. (2020). "Improved code summarization via a graph neural network". *arXiv preprint arXiv:2004.02843*.

LeCun, Y. and Y. Bengio. (1998). "Convolutional networks for images, speech, and time series". In: *The handbook of brain theory and neural networks*. 255–258.

Lee, D., C. Szegedy, M. Rabe, S. Loos, and K. Bansal. (2020). "Mathematical Reasoning in Latent Space". In: *International Conference on Learning Representations*. URL: https://openreview.net/forum?id=Ske31kBtPr.

Lee, H., Y. Peirsman, A. Chang, N. Chambers, M. Surdeanu, and D. Jurafsky. (2011). "Stanford's multi-pass sieve coreference resolution system at the conll-2011 shared task". In: *Proceedings of the 15th conference on computational natural language learning: Shared task*. Association for Computational Linguistics. 28–34.

Levesque, H., E. Davis, and L. Morgenstern. (2012). "The winograd schema challenge". In: *Thirteenth International Conference on the Principles of Knowledge Representation and Reasoning*. Citeseer.

Levi, F. W. (1942). *Finite geometrical systems: six public lectues delivered in February, 1940, at the University of Calcutta*. University of Calcutta.

Lewis, D. D., Y. Yang, T. G. Rose, and F. Li. (2004). "Rcv1: A new benchmark collection for text categorization research". *Journal of machine learning research*. 5(Apr): 361–397.

Li, C. and D. Goldwasser. (2019). "Encoding Social Information with Graph Convolutional Networks forPolitical Perspective Detection in News Media". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 2594–2604. DOI: 10.18653/v1/P19-1247.

Li, C., Y. Cao, L. Hou, J. Shi, J. Li, and T.-S. Chua. (2019). "Semi-supervised entity alignment via joint knowledge embedding model and cross-graph model". 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)*. 2723–2732.

Li, L., A. Way, and Q. Liu. (2017). "Context-Aware Graph Segmentation for Graph-Based Translation". In: *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*. Valencia, Spain: Association for Computational Linguistics. 599–604. URL: https://www.aclweb.org/anthology/E17-2095.

Li, R., S. Wang, F. Zhu, and J. Huang. (2018a). "Adaptive graph convolutional neural networks". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32.

Li, S., L. Wu, S. Feng, F. Xu, F. Xu, and S. Zhong. (2020a). "Graph-to-Tree Neural Networks for Learning Structured Input-Output Translation with Applications to Semantic Parsing and Math Word Problem". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 2841–2852. DOI: 10.18653/v1/2020.findings-emnlp.255.

Li, W., X. Xiao, J. Liu, H. Wu, H. Wang, and J. Du. (2020b). "Leveraging Graph to Improve Abstractive Multi-Document Summarization". *arXiv preprint arXiv:2005.10043*.

Li, Y., J. Amelot, X. Zhou, S. Bengio, and S. Si. (2020c). "Auto completion of user interface layout design using transformer-based tree decoders". *arXiv preprint arXiv:2001.05308*.

Li, Y., N. Duan, B. Zhou, X. Chu, W. Ouyang, X. Wang, and M. Zhou. (2018b). "Visual question generation as dual task of visual question answering". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 6116–6124.

Li, Y., D. Tarlow, M. Brockschmidt, and R. Zemel. (2015). "Gated graph sequence neural networks". *arXiv preprint arXiv:1511.05493*.

Li, Z., S. He, J. Cai, Z. Zhang, H. Zhao, G. Liu, L. Li, and L. Si. (2018c). "A unified syntax-aware framework for semantic role labeling". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 2401–2411.

Liao, K., L. Lebanoff, and F. Liu. (2018). "Abstract meaning representation for multi-document summarization". *arXiv preprint arXiv:1806.05655*.

Lin, B. Y., X. Chen, J. Chen, and X. Ren. (2019a). "KagNet: Knowledge-Aware Graph Networks for Commonsense Reasoning". 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)*. Hong Kong, China: Association for Computational Linguistics. 2829–2839. DOI: 10.18653/v1/D19-1282.

Lin, B. Y., X. Chen, J. Chen, and X. Ren. (2019b). "KagNet: Knowledge-Aware Graph Networks for Commonsense Reasoning". 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)*. 2822–2832.

Lin, X. V., R. Socher, and C. Xiong. (2018). "Multi-Hop Knowledge Graph Reasoning with Reward Shaping". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 3243–3253.

Linmei, H., T. Yang, C. Shi, H. Ji, and X. Li. (2019). "Heterogeneous Graph Attention Networks for Semi-supervised Short Text Classification". 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)*. Hong Kong, China: Association for Computational Linguistics. 4821–4830. DOI: 10.18653/v1/D19-1488.

Liu, B., D. Niu, H. Wei, J. Lin, Y. He, K. Lai, and Y. Xu. (2019a). "Matching Article Pairs with Graphical Decomposition and Convolutions". In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28-August 2, 2019, Volume 1: Long Papers*. Ed. by A. Korhonen, D. R. Traum, and L. Màrquez. Association for Computational Linguistics. 6284–6294.

Liu, B. and L. Wu. (2022). "Graph Neural Networks in Natural Language Processing". In: *Graph Neural Networks: Foundations, Frontiers, and Applications*. Ed. by L. Wu, P. Cui, J. Pei, and L. Zhao. Singapore: Springer Singapore. 463–481.

Liu, B., M. Zhao, D. Niu, K. Lai, Y. He, H. Wei, and Y. Xu. (2019b). "Learning to Generate Questions by LearningWhat Not to Generate". In: *The World Wide Web Conference. WWW '19*. San Francisco, CA, USA: Association for Computing Machinery. 1106–1118.

Liu, B., B. Ramsundar, P. Kawthekar, J. Shi, J. Gomes, Q. Luu Nguyen, S. Ho, J. Sloane, P. Wender, and V. Pande. (2017). "Retrosynthetic reaction prediction using neural sequence-to-sequence models". *ACS central science*. 3(10): 1103–1113.

Liu, D. and D. Gildea. (2010). "Semantic role features for machine translation". In: *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)*. 716–724.

Liu, J. and Y. Zhang. (2017). "In-Order Transition-based Constituent Parsing". *Transactions of the Association for Computational Linguistics*. 5: 413–424. DOI: 10.1162/tacl_a_00070.

Liu, M., Y. Luo, L. Wang, Y. Xie, H. Yuan, S. Gui, H. Yu, Z. Xu, J. Zhang, Y. Liu, *et al.* (2021a). "DIG: A Turnkey Library for Diving into Graph Deep Learning Research". *arXiv preprint arXiv:2103.12608*.

Liu, P., S. Chang, X. Huang, J. Tang, and J. C. K. Cheung. (2019c). "Contextualized non-local neural networks for sequence learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 6762–6769.

Liu, P., X. Qiu, and X. Huang. (2016). "Recurrent neural network for text classification with multi-task learning". In: *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*. 2873–2879.

Liu, P. J., M. Saleh, E. Pot, B. Goodrich, R. Sepassi, L. Kaiser, and N. Shazeer. (2018a). "Generating wikipedia by summarizing long sequences". *arXiv preprint arXiv:1801.10198*.

Liu, S., Y. Chen, X. Xie, J. K. Siow, and Y. Liu. (2021b). "Retrieval-Augmented Generation for Code Summarization via Hybrid GNN". In: *9th International Conference on Learning Representations*.

Liu, X., Z. Luo, and H. Huang. (2018b). "Jointly Multiple Events Extraction via Attention-based Graph Information Aggregation". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 1247–1256. DOI: 10.18653/v1/D18-1156.

Liu, X., X. You, X. Zhang, J. Wu, and P. Lv. (2020). "Tensor Graph Convolutional Networks for Text Classification". In: *The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, The Tenth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2020, New York, NY, USA, February 7-12, 2020*. AAAI Press. 8409–8416.

Liu, X., Q. Chen, C. Deng, H. Zeng, J. Chen, D. Li, and B. Tang. (2018c). "Lcqmc: A large-scale chinese question matching corpus". In: *Proceedings of the 27th International Conference on Computational Linguistics*. 1952–1962.

Liu, Y. (2019). "Fine-tune BERT for extractive summarization". *arXiv preprint arXiv:1903.10318*.

Liu, Y., M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. (2019d). "Roberta: A robustly optimized bert pretraining approach". *arXiv preprint arXiv:1907.11692*.

Liu, Y., S. Feng, D. Wang, K. Song, F. Ren, and Y. Zhang. (2021c). "A Graph Reasoning Network for Multi-turn Response Selection via Customized Pre-training". In: *The Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Virtual, February 2-9, 2021*.

Long, Q., Y. Jin, G. Song, Y. Li, and W. Lin. (2020). "Graph Structural-topic Neural Network". In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. KDD '20*. New York, NY, USA: Association for Computing Machinery. 1065–1073. DOI: 10.1145/3394486.3403150. (Accessed on 12/25/2020).

Lowe, R., N. Pow, I. V. Serban, and J. Pineau. (2015). "The Ubuntu Dialogue Corpus: A Large Dataset for Research in Unstructured Multi-Turn Dialogue Systems". In: *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 285.

Lu, Z. and H. Li. (2013). "A deep architecture for matching short texts". *Advances in neural information processing systems*. 26: 1367–1375.

Luan, Y., L. He, M. Ostendorf, and H. Hajishirzi. (2018). "Multi-Task Identification of Entities, Relations, and Coreference for Scientific Knowledge Graph Construction". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 3219–3232.

Luan, Y., D. Wadden, L. He, A. Shah, M. Ostendorf, and H. Hajishirzi. (2019). "A general framework for information extraction using dynamic span graphs". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 3036–3046.

Luke, S. (2005). "Ze lemoyer and Michael Collins. Learning to map sentences to logical form: Structured classification with probabilistic categorial grammars". In: *UAI*. Vol. 2. 3.

Luo, Y. and H. Zhao. (2020). "Bipartite Flat-Graph Network for Nested Named Entity Recognition". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 6408–6418. DOI: 10.18653/v1/2020.acl-main.571.

Luong, M.-T., H. Pham, and C. D. Manning. (2015). "Effective approaches to attention-based neural machine translation". *arXiv preprint arXiv:1508. 04025*.

Ma, X., Z. Hu, J. Liu, N. Peng, G. Neubig, and E. Hovy. (2018). "Stack-Pointer Networks for Dependency Parsing". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia: Association for Computational Linguistics. 1403–1414. DOI: 10.18653/v1/P18-1130.

Ma, Y., S. Wang, C. C. Aggarwal, and J. Tang. (2019). "Graph convolutional networks with eigenpooling". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 723–731.

Malaviya, C., C. Bhagavatula, A. Bosselut, and Y. Choi. (2020). "Commonsense Knowledge Base Completion with Structural and Semantic Context." In: *AAAI*. 2925–2933.

Mann, W. C. and S. A. Thompson. (1987). *Rhetorical structure theory: A theory of text organization*. University of Southern California, Information Sciences Institute Los Angeles.

Marcheggiani, D., J. Bastings, and I. Titov. (2018). "Exploiting Semantics in Neural Machine Translation with Graph Convolutional Networks". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. New Orleans, Louisiana: Association for Computational Linguistics. 486–492. DOI: 10.18653/v1/N18-2078.

Marcheggiani, D., A. Frolov, and I. Titov. (2017). "A simple and accurate syntax-agnostic neural model for dependency-based semantic role labeling". *arXiv preprint arXiv:1701.02593*.

Marcheggiani, D. and L. Perez-Beltrachini. (2018). "Deep Graph Convolutional Encoders for Structured Data to Text Generation". In: *Proceedings of the 11th International Conference on Natural Language Generation*. Tilburg University, The Netherlands: Association for Computational Linguistics. 1–9. DOI: 10.18653/v1/W18-6501.

Marcheggiani, D. and I. Titov. (2017). "Encoding Sentences with Graph Convolutional Networks for Semantic Role Labeling". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 1506–1515.

Marcheggiani, D. and I. Titov. (2020). "Graph Convolutions over Constituent Trees for Syntax-Aware Semantic Role Labeling". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 3915–3928.

Marquez, L., P. Comas, J. Giménez, and N. Catala. (2005). "Semantic role labeling as sequential tagging". In: *Proceedings of the Ninth Conference on Computational Natural Language Learning (CoNLL-2005)*. 193–196.

Miao, Y., L. Yu, and P. Blunsom. (2016). "Neural variational inference for text processing". In: *International conference on machine learning*. PMLR. 1727–1736.

Mihalcea, R. (2005). "Unsupervised large-vocabulary word sense disambiguation with graph-based algorithms for sequence data labeling". In: *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*. 411–418.

Mihalcea, R. and D. Radev. (2011). *Graph-based natural language processing and information retrieval*. Cambridge university press.

Mihalcea, R. and P. Tarau. (2004). "Textrank: Bringing order into text". In: *Proceedings of the 2004 conference on empirical methods in natural language processing*. 404–411.

Mihaylov, T., P. Clark, T. Khot, and A. Sabharwal. (2018). "Can a Suit of Armor Conduct Electricity? A New Dataset for Open Book Question Answering". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 2381–2391.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. (2013). "Distributed Representations of Words and Phrases and their Compositionality". In: *Advances in Neural Information Processing Systems 26: 27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*. Ed. by C. J. C. Burges, L. Bottou, Z. Ghahramani, and K. Q. Weinberger. 3111–3119. URL: https://proceedings.neurips.cc/paper/2013/hash/9aa42b31882ec039965f3c4923ce901b-Abstract.html.

Miller, A., A. Fisch, J. Dodge, A.-H. Karimi, A. Bordes, and J. Weston. (2016). "Key-Value Memory Networks for Directly Reading Documents". In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*. Austin, Texas: Association for Computational Linguistics. 1400–1409. DOI: 10.18653/v1/D16-1147.

Minkov, E., W. W. Cohen, and A. Y. Ng. (2006). "Contextual search and name disambiguation in email using graphs". In: *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*. 27–34.

Monz, C. and B. J. Dorr. (2005). "Iterative translation disambiguation for cross-language information retrieval". In: *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*. 520–527.

Nair, V. and G. E. Hinton. (2010). "Rectified linear units improve restricted boltzmann machines". In: *ICML*.

Nallapati, R., B. Zhou, C. dos Santos, Ç. Gülçehre, and B. Xiang. (2016). "Abstractive Text Summarization using Sequence-to-sequence RNNs and Beyond". In: *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*. 280–290.

Nathani, D., J. Chauhan, C. Sharma, and M. Kaul. (2019a). "Learning Attention-based Embeddings for Relation Prediction in Knowledge Graphs". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 4710–4723. DOI: 10.18653/v1/P19-1466.

Nathani, D., J. Chauhan, C. Sharma, and M. Kaul. (2019b). "Learning Attention-based Embeddings for Relation Prediction in Knowledge Graphs". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 4710–4723.

Nguyen, T. H. and R. Grishman. (2018). "Graph convolutional networks with argument-aware pooling for event detection". In: *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*. AAAI press. 5900–5907.

Niu, Z.-Y., D. Ji, and C. L. Tan. (2005). "Word sense disambiguation using label propagation based semi-supervised learning". In: *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05)*. 395–402.

Nivre et al., J. (2018). "Universal Dependencies 2.2". URL: http://hdl.handle.net/11234/1-2837.

Norcliffe-Brown, W., S. Vafeias, and S. Parisot. (2018). "Learning Conditioned Graph Structures for Interpretable Visual Question Answering". In: *Advances in neural information processing systems*. 8334–8343.

Page, L., S. Brin, R. Motwani, and T. Winograd. (1999). "The PageRank citation ranking: Bringing order to the web." *Tech. rep.* Stanford InfoLab.

Palangi, H., L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song, and R. Ward. (2016). "Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval". *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 24(4): 694–707.

Pan, L., Y. Xie, Y. Feng, T.-S. Chua, and M.-Y. Kan. (2020). "Semantic Graphs for Generating Deep Questions". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 1463–1475.

Pang, B. and L. Lee. (2004). "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts". In: *Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL-04)*. 271–278.

Pang, B., L. Lee, and S. Vaithyanathan. (2002). "Thumbs up? Sentiment Classification using Machine Learning Techniques". In: *EMNLP*.

Pang, L., Y. Lan, J. Guo, J. Xu, S. Wan, and X. Cheng. (2016). "Text matching as image recognition". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. No. 1.

Paulus, R., C. Xiong, and R. Socher. (2018). "A Deep Reinforced Model for Abstractive Summarization". In: *International Conference on Learning Representations*.

Peng, H., J. Li, Y. He, Y. Liu, M. Bao, L. Wang, Y. Song, and Q. Yang. (2018). "Large-scale hierarchical text classification with recursively regularized deep graph-cnn". In: *Proceedings of the 2018 world wide web conference*. 1063–1072.

Peng, H., N. Pappas, D. Yogatama, R. Schwartz, N. Smith, and L. Kong. (2021). "Random Feature Attention". In: *International Conference on Learning Representations*.

Pennington, J., R. Socher, and C. D. Manning. (2014). "Glove: Global vectors for word representation". In: *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. 1532–1543.

Peters, M., M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. (2018). "Deep Contextualized Word Representations". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. 2227–2237.

Phan, X.-H., L.-M. Nguyen, and S. Horiguchi. (2008). "Learning to classify short and sparse text & web with hidden topics from large-scale data collections". In: *Proceedings of the 17th international conference on World Wide Web*. 91–100.

Ponte, J. M. and W. B. Croft. (1998). "A language modeling approach to information retrieval". In: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. 275–281.

Pontiki, M., D. Galanis, H. Papageorgiou, I. Androutsopoulos, S. Manandhar, M. Al-Smadi, M. Al-Ayyoub, Y. Zhao, B. Qin, O. De Clercq, *et al.* (2016). "Semeval-2016 task 5: Aspect based sentiment analysis". In: *International workshop on semantic evaluation*. 19–30.

Pontiki, M., D. Galanis, H. Papageorgiou, S. Manandhar, and I. Androutsopoulos. (2015). "Semeval-2015 task 12: Aspect based sentiment analysis". In: *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*. 486–495.

Pontiki, M., D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar. (2014). "SemEval-2014 Task 4: Aspect Based Sentiment Analysis". In: *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*. Dublin, Ireland: Association for Computational Linguistics. 27–35. DOI: 10.3115/v1/S14-2004.

Pouran Ben Veyseh, A., N. Nouri, F. Dernoncourt, Q. H. Tran, D. Dou, and T. H. Nguyen. (2020). "Improving Aspect-based Sentiment Analysis with Gated Graph Convolutional Networks and Syntax-based Regulation". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 4543–4548. DOI: 10.18653/v1/2020.findings-emnlp.407. (Accessed on 12/26/2020).

Pourdamghani, N., Y. Gao, U. Hermjakob, and K. Knight. (2014). "Aligning English Strings with Abstract Meaning Representation Graphs". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics. 425–429. DOI: 10.3115/v1/D14-1048.

Pourdamghani, N., K. Knight, and U. Hermjakob. (2016). "Generating english from abstract meaning representations". In: *Proceedings of the 9th international natural language generation conference*. 21–25.

Qian, Y., E. Santus, Z. Jin, J. Guo, and R. Barzilay. (2019). "GraphIE: A Graph-Based Framework for Information Extraction". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 751–761.

Qiu, J., Q. Chen, Y. Dong, J. Zhang, H. Yang, M. Ding, K. Wang, and J. Tang. (2020). "Gcc: Graph contrastive coding for graph neural network pre-training". In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1150–1160.

Qiu, L., Y. Xiao, Y. Qu, H. Zhou, L. Li, W. Zhang, and Y. Yu. (2019). "Dynamically Fused Graph Network for Multi-hop Reasoning". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 6140–6150. DOI: 10.18653/v1/P19-1617.

Qu, M., T. Gao, L.-P. Xhonneux, and J. Tang. (2020). "Few-shot Relation Extraction via Bayesian Meta-learning on Relation Graphs". In: *International Conference on Machine Learning*. PMLR. 7867–7876.

Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever. (2019). "Language models are unsupervised multitask learners". *OpenAI blog*. 1(8): 9.

Rajpurkar, P., J. Zhang, K. Lopyrev, and P. Liang. (2016). "SQuAD: 100, 000+ Questions for Machine Comprehension of Text". In: *EMNLP*.

Ramage, D., A. N. Rafferty, and C. D. Manning. (2009). "Random walks for text semantic similarity". In: *Proceedings of the 2009 workshop on graph-based methods for natural language processing (TextGraphs-4)*. 23–31.

Ran, Q., Y. Lin, P. Li, J. Zhou, and Z. Liu. (2019). "NumNet: Machine Reading Comprehension with Numerical Reasoning". 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 2019, Hong Kong, China, November 3-7, 2019*. Ed. by K. Inui, J. Jiang, V. Ng, and X. Wan. Association for Computational Linguistics. 2474–2484.

Rashkin, H., M. Sap, E. Allaway, N. A. Smith, and Y. Choi. (2018). "Event2mind: Commonsense inference on events, intents, and reactions". *arXiv preprint arXiv:1805.06939*.

Reddy, S., D. Chen, and C. D. Manning. (2019). "Coqa: A conversational question answering challenge". *Transactions of the Association for Computational Linguistics*. 7: 249–266.

Redmon, J. and A. Farhadi. (2018). "Yolov3: An incremental improvement". *arXiv preprint arXiv:1804.02767*.

Ribeiro, L. F. R., C. Gardent, and I. Gurevych. (2019a). "Enhancing AMR-to-Text Generation with Dual Graph Representations". 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)*. Hong Kong, China: Association for Computational Linguistics. 3183–3194. DOI: 10.18653/v1/D19-1314.

Ribeiro, L. F., C. Gardent, and I. Gurevych. (2019b). "Enhancing amr-to-text generation with dual graph representations". *arXiv preprint arXiv:1909. 00352*.

Riedel, S., L. Yao, and A. McCallum. (2010). "Modeling relations and their mentions without labeled text". In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer. 148–163.

Sachan, D. S., L. Wu, M. Sachan, and W. Hamilton. (2020). "Stronger Transformers for Neural Multi-Hop Question Generation". *arXiv preprint arXiv: 2010.11374*.

Sagae, K. and A. Lavie. (2005). "A Classifier-Based Parser with Linear Run-Time Complexity". In: *Proceedings of the Ninth International Workshop on Parsing Technology*. Vancouver, British Columbia: Association for Computational Linguistics. 125–132. URL: https://www.aclweb.org/anthology/W05-1513.

Sahu, S. K., F. Christopoulou, M. Miwa, and S. Ananiadou. (2019). "Inter-sentence Relation Extraction with Document-level Graph Convolutional Neural Network". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 4309–4316. DOI: 10.18653/v1/P19-1423.

Sandhaus, E. (2008). "The New York Times Annotated Corpus". Version V1. DOI: 11272.1/AB2/GZC6PL.

Santoro, A., D. Raposo, D. G. Barrett, M. Malinowski, R. Pascanu, P. Battaglia, and T. Lillicrap. (2017). "A simple neural network module for relational reasoning". In: *Proceedings of the 31st International Conference on Neural Information Processing Systems*. 4974–4983.

Saxena, A., A. Tripathi, and P. Talukdar. (2020). "Improving Multi-hop Question Answering over Knowledge Graphs using Knowledge Base Embeddings". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 4498–4507. DOI: 10.18653/v1/2020.acl-main.412.

Schlichtkrull, M., T. N. Kipf, P. Bloem, R. Van Den Berg, I. Titov, and M. Welling. (2018). "Modeling relational data with graph convolutional networks". In: *European Semantic Web Conference*. Springer. 593–607.

Schuster, M. and K. K. Paliwal. (1997). "Bidirectional recurrent neural networks". *IEEE Transactions on Signal Processing*. 45(11): 2673–2681. DOI: 10.1109/78.650093.

See, A., P. J. Liu, and C. D. Manning. (2017). "Get To The Point: Summarization with Pointer-Generator Networks". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 1073–1083.

Seo, M. J., A. Kembhavi, A. Farhadi, and H. Hajishirzi. (2017). "Bidirectional Attention Flow for Machine Comprehension". In: *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net. URL: https://openreview.net/forum?id=HJ0UKP9ge.

Serban, I., A. Sordoni, Y. Bengio, A. Courville, and J. Pineau. (2016). "Building end-to-end dialogue systems using generative hierarchical neural network models". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. No. 1.

Serban, I., A. Sordoni, R. Lowe, L. Charlin, J. Pineau, A. Courville, and Y. Bengio. (2017). "A hierarchical latent variable encoder-decoder model for generating dialogues". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. No. 1.

Shang, C., Y. Tang, J. Huang, J. Bi, X. He, and B. Zhou. (2019). "End-to-end structure-aware convolutional networks for knowledge base completion". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 3060–3067.

Shao, B., Y. Gong, W. Qi, G. Cao, J. Ji, and X. Lin. (2020). "Graph-Based Transformer with Cross-Candidate Verification for Semantic Parsing". *Proceedings of the AAAI Conference on Artificial Intelligence*. 34(05): 8807–8814. DOI: 10.1609/aaai.v34i05.6408.

Shaw, P., J. Uszkoreit, and A. Vaswani. (2018). "Self-Attention with Relative Position Representations". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. New Orleans, Louisiana: Association for Computational Linguistics. 464–468. DOI: 10.18653/v1/N18-2074.

Shen, D. and M. Lapata. (2007). "Using semantic roles to improve question answering". In: *Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL)*. 12–21.

Shen, Z., M. Zhang, H. Zhao, S. Yi, and H. Li. (2021). "Efficient attention: Attention with linear complexities". In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 3531–3539.

Shi, C., M. Xu, H. Guo, M. Zhang, and J. Tang. (2020). "A graph to graphs framework for retrosynthesis prediction". In: *International Conference on Machine Learning*. PMLR. 8818–8827.

Simonovsky, M. and N. Komodakis. (2017). "Dynamic edge-conditioned filters in convolutional neural networks on graphs". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 3693–3702.

Song, L., D. Gildea, Y. Zhang, Z. Wang, and J. Su. (2019). "Semantic neural machine translation using AMR". *Transactions of the Association for Computational Linguistics*. 7: 19–31.

Song, L., X. Peng, Y. Zhang, Z. Wang, and D. Gildea. (2017). "AMR-to-text Generation with Synchronous Node Replacement Grammar". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 7–13.

Song, L., A. Wang, J. Su, Y. Zhang, K. Xu, Y. Ge, and D. Yu. (2020). "Structural Information Preserving for Graph-to-Text Generation". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 7987–7998. DOI: 10.18653/v1/2020.acl-main.712.

Song, L., Z. Wang, W. Hamza, Y. Zhang, and D. Gildea. (2018a). "Leveraging context information for natural question generation". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. 569–574.

Song, L., Z. Wang, M. Yu, Y. Zhang, R. Florian, and D. Gildea. (2018b). "Exploring graph-structured passage representation for multi-hop reading comprehension with graph neural networks". *arXiv preprint arXiv:1809.02040*.

Song, L., Y. Zhang, Z. Wang, and D. Gildea. (2018c). "A Graph-to-Sequence Model for AMR-to-Text Generation". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Melbourne, Australia: Association for Computational Linguistics. 1616–1626. DOI: 10.18653/v1/P18-1150.

Song, L., Y. Zhang, Z. Wang, and D. Gildea. (2018d). "A graph-to-sequence model for amr-to-text generation". *arXiv preprint arXiv:1805.02473*.

Song, L., Y. Zhang, Z. Wang, and D. Gildea. (2018e). "N-ary Relation Extraction using Graph-State LSTM". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 2226–2235. DOI: 10.18653/v1/D18-1246.

Sorokin, D. and I. Gurevych. (2018a). "Modeling Semantics with Gated Graph Neural Networks for Knowledge Base Question Answering". In: *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, New Mexico, USA, August 20-26, 2018*. Ed. by E. M. Bender, L. Derczynski, and P. Isabelle. Association for Computational Linguistics. 3306–3317.

Sorokin, D. and I. Gurevych. (2018b). "Modeling semantics with gated graph neural networks for knowledge base question answering". *arXiv preprint arXiv:1808.04126*.

Speer, R., J. Chin, and C. Havasi. (2017). "Conceptnet 5.5: An open multilingual graph of general knowledge". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 31. No. 1.

Su, D., Y. Xu, W. Dai, Z. Ji, T. Yu, and P. Fung. (2020). "Multi-hop Question Generation with Graph Convolutional Network". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 4636–4647. DOI: 10.18653/v1/2020.findings-emnlp.416.

Suchanek, F. M., G. Kasneci, and G. Weikum. (2008). "Yago: A large ontology from wikipedia and wordnet". *Journal of Web Semantics*. 6(3): 203–217.

Sui, D., Y. Chen, K. Liu, J. Zhao, and S. Liu. (2019). "Leverage lexical knowledge for chinese named entity recognition via collaborative graph network". 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)*. 3821–3831.

Sukhbaatar, S., A. Szlam, J. Weston, and R. Fergus. (2015). "End-to-end memory networks". In: *Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 2*. 2440–2448.

Sun, C., Y. Gong, Y. Wu, M. Gong, D. Jiang, M. Lan, S. Sun, and N. Duan. (2019a). "Joint Type Inference on Entities and Relations via Graph Convolutional Networks". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 1361–1370. DOI: 10.18653/v1/P19-1131.

Sun, H., T. Bedrax-Weiss, and W. Cohen. (2019b). "PullNet: Open Domain Question Answering with Iterative Retrieval on Knowledge Bases and Text". 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)*. Hong Kong, China: Association for Computational Linguistics. 2380–2390. DOI: 10.18653/v1/D19-1242.

Sun, H., B. Dhingra, M. Zaheer, K. Mazaitis, R. Salakhutdinov, and W. Cohen. (2018a). "Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 4231–4242. DOI: 10.18653/v1/D18-1455.

Sun, K., R. Zhang, S. Mensah, Y. Mao, and X. Liu. (2019c). "Aspect-Level Sentiment Analysis Via Convolution over Dependency Tree". 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)*. Hong Kong, China: Association for Computational Linguistics. 5679–5688. DOI: 10.18653/v1/D19-1569.

Sun, T., Y. Shao, X. Qiu, Q. Guo, Y. Hu, X. Huang, and Z. Zhang. (2020a). "CoLAKE: Contextualized Language and Knowledge Embedding". In: *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*. International Committee on Computational Linguistics. 3660–3670.

Sun, Y., J. Han, X. Yan, P. S. Yu, and T. Wu. (2011). "Pathsim: Meta path-based top-k similarity search in heterogeneous information networks". *Proceedings of the VLDB Endowment*. 4(11): 992–1003.

Sun, Z., W. Hu, and C. Li. (2017). "Cross-lingual entity alignment via joint attribute-preserving embedding". In: *International Semantic Web Conference*. Springer. 628–644.

Sun, Z., W. Hu, Q. Zhang, and Y. Qu. (2018b). "Bootstrapping Entity Alignment with Knowledge Graph Embedding." In: *IJCAI*. Vol. 18. 4396–4402.

Sun, Z., C. Wang, W. Hu, M. Chen, J. Dai, W. Zhang, and Y. Qu. (2020b). "Knowledge graph alignment network with gated multi-hop neighborhood aggregation". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 01. 222–229.

Sun, Z., Q. Zhu, Y. Xiong, Y. Sun, L. Mou, and L. Zhang. (2020c). "Treegen: A tree-based transformer architecture for code generation". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 05. 8984–8991.

Sutskever, I., O. Vinyals, and Q. V. Le. (2014). "Sequence to Sequence Learning with Neural Networks". In: *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger. 3104–3112. URL: https://proceedings.neurips.cc/paper/2014/hash/a14ac55a4f27472c5d894ec1c3c743d2-Abstract.html.

Tai, K. S., R. Socher, and C. D. Manning. (2015). "Improved Semantic Representations From Tree-Structured Long Short-Term Memory Networks". In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 1556–1566.

Talmor, A. and J. 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)*. 641–651.

Talmor, A., J. Herzig, N. Lourie, and J. Berant. (2019). "CommonsenseQA: A Question Answering Challenge Targeting Commonsense Knowledge". In: *NAACL-HLT (1)*.

Tang, D., B. Qin, X. Feng, and T. Liu. (2016). “Effective LSTMs for Target-Dependent Sentiment Classification”. In: *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*. 3298–3307.

Tang, H., D. Ji, C. Li, and Q. Zhou. (2020a). “Dependency Graph Enhanced Dual-transformer Structure for Aspect-based Sentiment Classification”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 6578–6588. DOI: 10.18653/v1/2020.acl-main.588.

Tang, J., M. Qu, and Q. Mei. (2015). “Pte: Predictive text embedding through large-scale heterogeneous text networks”. In: *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 1165–1174.

Tang, Z., Y. Shen, X. Ma, W. Xu, J. Yu, and W. Lu. (2020b). “Multi-hop Reading Comprehension across Documents with Path-based Graph Convolutional Network”. In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI 2020*. Ed. by C. Bessiere. ijcai.org. 3905–3911.

Tang, Z., Y. Shen, X. Ma, W. Xu, J. Yu, and W. Lu. (2020c). “Multi-hop reading comprehension across documents with path-based graph convolutional network”. *arXiv preprint arXiv:2006.06478*.

Tarau, P., R. Mihalcea, and E. Figa. (2005). “Semantic document engineering with WordNet and PageRank”. In: *Proceedings of the 2005 ACM symposium on Applied computing*. 782–786.

Taylor, A., M. Marcus, and B. Santorini. (2003). “The Penn treebank: an overview”. *Treebanks*: 5–22.

Teru, K., E. Denis, and W. Hamilton. (2020). “Inductive relation prediction by subgraph reasoning”. In: *International Conference on Machine Learning*. PMLR. 9448–9457.

Thayaparan, M., M. Valentino, V. Schlegel, and A. Freitas. (2019). “Identifying Supporting Facts for Multi-hop Question Answering with Document Graph Networks”. In: *Proceedings of the Thirteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-13)*. Hong Kong: Association for Computational Linguistics. 42–51. DOI: 10.18653/v1/D19-5306.

Thompson, A. (2017). “All the news: 143,000 articles from 15 american publications”.

Toutanova, K., D. Chen, P. Pantel, H. Poon, P. Choudhury, and M. Gamon. (2015). “Representing text for joint embedding of text and knowledge bases”. In: *Proceedings of the 2015 conference on empirical methods in natural language processing*. 1499–1509.

Trischler, A., T. Wang, X. Yuan, J. Harris, A. Sordoni, P. Bachman, and K. Suleman. (2017). “NewsQA: A Machine Comprehension Dataset”. In: *Proceedings of the 2nd Workshop on Representation Learning for NLP*. 191–200.

Trouillon, T., J. Welbl, S. Riedel, É. Gaussier, and G. Bouchard. (2016). “Complex embeddings for simple link prediction”. In: *International Conference on Machine Learning*. PMLR. 2071–2080.

Tsai, Y.-H. H., S. Bai, M. Yamada, L.-P. Morency, and R. Salakhutdinov. (2019). “Transformer Dissection: An Unified Understanding for Transformer’s Attention via the Lens of Kernel”. 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)*. 4335–4344.

Tu, M., G. Wang, J. Huang, Y. Tang, X. He, and B. Zhou. (2019a). “Multi-hop Reading Comprehension across Multiple Documents by Reasoning over Heterogeneous Graphs”. In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 2704–2713. DOI: 10.18653/v1/P19-1260.

Tu, M., G. Wang, J. Huang, Y. Tang, X. He, and B. Zhou. (2019b). “Multi-hop Reading Comprehension across Multiple Documents by Reasoning over Heterogeneous Graphs”. In: *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*. Ed. by A. Korhonen, D. R. Traum, and L. Màrquez. Association for Computational Linguistics. 2704–2713.

Tu, Z., Z. Lu, Y. Liu, X. Liu, and H. Li. (2016). “Modeling coverage for neural machine translation”. *arXiv preprint arXiv:1601.04811*.

Usbeck, R., A.-C. N. Ngomo, B. Haarmann, A. Krithara, M. Röder, and G. Napolitano. (2017). "7th open challenge on question answering over linked data (QALD-7)". In: *Semantic web evaluation challenge*. Springer. 59–69.

Vashishth, S., R. Joshi, S. S. Prayaga, C. Bhattacharyya, and P. Talukdar. (2018). "RESIDE: Improving Distantly-Supervised Neural Relation Extraction using Side Information". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 1257–1266. DOI: 10.18653/v1/D18-1157.

Vashishth, S., S. Sanyal, V. Nitin, and P. Talukdar. (2019). "Composition-based multi-relational graph convolutional networks". *arXiv preprint arXiv:1911.03082*.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. (2017). "Attention is All you Need". In: *NIPS*.

Velickovic, P., L. Buesing, M. C. Overlan, R. Pascanu, O. Vinyals, and C. Blundell. (2020). "Pointer Graph Networks". In: *Advances in Neural Information Processing Systems*.

Velickovic, P., G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio. (2018). "Graph Attention Networks". In: *6th International Conference on Learning Representations*.

Vinyals, O., S. Bengio, and M. Kudlur. (2016). "Order Matters: Sequence to sequence for sets". In: *4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings*. Ed. by Y. Bengio and Y. LeCun. URL: http://arxiv.org/abs/1511.06391.

Vinyals, O., M. Fortunato, and N. Jaitly. (2015). "Pointer networks". In: *Advances in Neural Information Processing Systems*. 2692–2700.

Vitale, D., P. Ferragina, and U. Scaiella. (2012). "Classification of short texts by deploying topical annotations". In: *European Conference on Information Retrieval*. Springer. 376–387.

Wan, S., Y. Lan, J. Guo, J. Xu, L. Pang, and X. Cheng. (2016). "A deep architecture for semantic matching with multiple positional sentence representations". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 30. No. 1.

Wang, D., P. Liu, Y. Zheng, X. Qiu, and X. Huang. (2020a). "Heterogeneous Graph Neural Networks for Extractive Document Summarization". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 6209–6219. DOI: 10.18653/v1/2020.acl-main.553.

Wang, H., S. Li, R. Pan, and M. Mao. (2019a). "Incorporating Graph Attention Mechanism into Knowledge Graph Reasoning Based on Deep Reinforcement Learning". 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)*. Hong Kong, China: Association for Computational Linguistics. 2623–2631. DOI: 10.18653/v1/D19-1264.

Wang, K., W. Shen, Y. Yang, X. Quan, and R. Wang. (2020b). "Relational Graph Attention Network for Aspect-based Sentiment Analysis". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 3229–3238. DOI: 10.18653/v1/2020.acl-main.295. (Accessed on 12/26/2020).

Wang, L., Z. Xu, Z. Lin, H. Zheng, and Y. Shen. (2020c). "Answer-driven Deep Question Generation based on Reinforcement Learning". In: *Proceedings of the 28th International Conference on Computational Linguistics*. Barcelona, Spain (Online): International Committee on Computational Linguistics. 5159–5170. DOI: 10.18653/v1/2020.coling-main.452.

Wang, L., Z. Xu, Z. Lin, H. Zheng, and Y. Shen. (2020d). "Answer-driven Deep Question Generation based on Reinforcement Learning". In: *Proceedings of the 28th International Conference on Computational Linguistics, COLING 2020, Barcelona, Spain (Online), December 8-13, 2020*. Ed. by D. Scott, N. Bel, and C. Zong. International Committee on Computational Linguistics. 5159–5170.

Wang, M., L. Yu, D. Zheng, Q. Gan, Y. Gai, Z. Ye, M. Li, J. Zhou, Q. Huang, C. Ma, *et al.* (2019b). "Deep Graph Library: Towards Efficient and Scalable Deep Learning on Graphs."

Wang, P., J. Han, C. Li, and R. Pan. (2019c). "Logic attention based neighborhood aggregation for inductive knowledge graph embedding". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 7152–7159.

Wang, R., D. Zhou, and Y. He. (2019d). "Atm: Adversarial-neural topic model". *Information Processing & Management*. 56(6): 102098.

Wang, S., B. Z. Li, M. Khabsa, H. Fang, and H. Ma. (2020e). "Linformer: Self-Attention with Linear Complexity". *CoRR*. abs/2006.04768. arXiv: 2006.04768. URL: https://arxiv.org/abs/2006.04768.

Wang, T., X. Wan, and H. Jin. (2020f). "AMR-To-Text Generation with Graph Transformer". *Transactions of the Association for Computational Linguistics*. 8: 19–33.

Wang, T., X. Wan, and S. Yao. (2020g). "Better AMR-To-Text Generation with Graph Structure Reconstruction". In: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Ed. by C. Bessiere. International Joint Conferences on Artificial Intelligence Organization. 3919–3925. DOI: 10.24963/ijcai.2020/542.

Wang, X., H. Ji, C. Shi, B. Wang, Y. Ye, P. Cui, and P. S. Yu. (2019e). "Heterogeneous graph attention network". In: *The World Wide Web Conference*. 2022–2032.

Wang, X., P. Kapanipathi, R. Musa, M. Yu, K. Talamadupula, I. Abdelaziz, M. Chang, A. Fokoue, B. Makni, N. Mattei, and M. Witbrock. (2019f). "Improving Natural Language Inference Using External Knowledge in the Science Questions Domain". *Proceedings of the AAAI Conference on Artificial Intelligence*. 33(01): 7208–7215. DOI: 10.1609/aaai.v33i01.33017208.

Wang, Y., X. Liu, and S. Shi. (2017). "Deep neural solver for math word problems". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 845–854.

Wang, Z., Q. Lv, X. Lan, and Y. Zhang. (2018). "Cross-lingual knowledge graph alignment via graph convolutional networks". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 349–357.

Wang, Z., J. Yang, and X. Ye. (2020h). "Knowledge Graph Alignment with Entity-Pair Embedding". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 1672–1680. DOI: 10.18653/v1/2020.emnlp-main.130.

Wang, Z., Z. Ren, C. He, P. Zhang, and Y. Hu. (2019g). "Robust Embedding with Multi-Level Structures for Link Prediction." In: *IJCAI*. 5240–5246.

Welbl, J., P. Stenetorp, and S. Riedel. (2018). "Constructing datasets for multi-hop reading comprehension across documents". *Transactions of the Association for Computational Linguistics*. 6: 287–302.

Williams, A., N. Nangia, and S. Bowman. (2018). "A Broad-Coverage Challenge Corpus for Sentence Understanding through Inference". *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. DOI: 10.18653/v1/n18-1101.

Williams, J. D., M. Henderson, A. Raux, B. Thomson, A. Black, and D. Ramachandran. (2014). "The dialog state tracking challenge series". *AI Magazine*. 35(4): 121–124.

Williams, R. J. (1992). "Simple statistical gradient-following algorithms for connectionist reinforcement learning". *Machine learning*. 8(3-4): 229–256.

Wu, J., M. Cao, J. C. K. Cheung, and W. L. Hamilton. (2020a). "TeMP: Temporal Message Passing for Temporal Knowledge Graph Completion". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 5730–5746. DOI: 10.18653/v1/2020.emnlp-main.462.

Wu, L., P. Cui, J. Pei, and L. Zhao. (2022). *Graph Neural Networks: Foundations, Frontiers, and Applications*. Singapore: Springer Singapore. 725.

Wu, Q., Q. Zhang, J. Fu, and X. Huang. (2020b). "A Knowledge-Aware Sequence-to-Tree Network for Math Word Problem Solving". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 7137–7146. DOI: 10.18653/v1/2020.emnlp-main.579.

Wu, Y., X. Liu, Y. Feng, Z. Wang, R. Yan, and D. Zhao. (2019a). "Relation-aware entity alignment for heterogeneous knowledge graphs". In: *Proceedings of the 28th International Joint Conference on Artificial Intelligence*. AAAI Press. 5278–5284.

Wu, Y., X. Liu, Y. Feng, Z. Wang, and D. Zhao. (2019b). "Jointly Learning Entity and Relation Representations for Entity Alignment". 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)*. 240–249.

Wu, Z., R. Koncel-Kedziorski, M. Ostendorf, and H. Hajishirzi. (2020c). "Extracting Summary Knowledge Graphs from Long Documents". *arXiv preprint arXiv:2009.09162*.

Xia, M., G. Huang, L. Liu, and S. Shi. (2019). "Graph Based Translation Memory for Neural Machine Translation". *Proceedings of the AAAI Conference on Artificial Intelligence*. 33(01): 7297–7304. DOI: 10.1609/aaai.v33i01.33017297.

Xia, Q., R. Wang, Z. Li, Y. Zhang, and M. Zhang. (2020). "Semantic Role Labeling with Heterogeneous Syntactic Knowledge". In: *Proceedings of the 28th International Conference on Computational Linguistics*. 2979–2990.

Xiao, F., J. Li, H. Zhao, R. Wang, and K. Chen. (2019). "Lattice-Based Transformer Encoder for Neural Machine Translation". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 3090–3097. DOI: 10.18653/v1/P19-1298.

Xie, Z., G. Zhou, J. Liu, and X. Huang. (2020). "ReInceptionE: Relation-aware inception network with joint local-global structural information for knowledge graph embedding". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 5929–5939.

Xiong, C., V. Zhong, and R. Socher. (2017a). "Dynamic Coattention Networks For Question Answering". In: *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net. URL: https://openreview.net/forum?id=rJeKjwvclx.

Xiong, W., T. Hoang, and W. Y. Wang. (2017b). "DeepPath: A Reinforcement Learning Method for Knowledge Graph Reasoning". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 564–573.

Xu, J., Z. Gan, Y. Cheng, and J. Liu. (2020a). "Discourse-aware neural extractive text summarization". In: *Proceedings of the 58th annual meeting of the association for computational linguistics*. 5021–5031.

Xu, K., L. Song, Y. Feng, Y. Song, and D. Yu. (2020b). "Coordinated Reasoning for Cross-Lingual Knowledge Graph Alignment". *arXiv preprint arXiv:2001.08728*.

Xu, K., L. Wang, M. Yu, Y. Feng, Y. Song, Z. Wang, and D. Yu. (2019a). "Cross-lingual Knowledge Graph Alignment via Graph Matching Neural Network". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 3156–3161. DOI: 10.18653/v1/P19-1304.

Xu, K., L. Wu, Z. Wang, Y. Feng, and V. Sheinin. (2018a). "SQL-to-Text Generation with Graph-to-Sequence Model". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 931–936. DOI: 10.18653/v1/D18-1112.

Xu, K., L. Wu, Z. Wang, Y. Feng, M. Witbrock, and V. Sheinin. (2018b). "Graph2seq: Graph to sequence learning with attention-based neural networks". *arXiv preprint arXiv:1804.00823*.

Xu, K., L. Wu, Z. Wang, M. Yu, L. Chen, and V. Sheinin. (2018c). "Exploiting Rich Syntactic Information for Semantic Parsing with Graph-to-Sequence Model". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 918–924. DOI: 10.18653/v1/D18-1110.

Xu, M., L. Li, D. F. Wai, Q. Liu, and L. S. Chao. (2020c). "Document Graph for Neural Machine Translation". *ArXiv*. abs/2012.03477.

Xu, X., W. Feng, Y. Jiang, X. Xie, Z. Sun, and Z.-H. Deng. (2019b). "Dynamically Pruned Message Passing Networks for Large-Scale Knowledge Graph Reasoning". *arXiv preprint arXiv:1909.11334*.

Yan, H., X. Jin, X. Meng, J. Guo, and X. Cheng. (2019). "Event Detection with Multi-Order Graph Convolution and Aggregated Attention". 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)*. Hong Kong, China: Association for Computational Linguistics. 5766–5770. DOI: 10.18653/v1/D19-1582.

Yang, B., W.-t. Yih, X. He, J. Gao, and L. Deng. (2014). "Embedding entities and relations for learning and inference in knowledge bases". *arXiv preprint arXiv:1412.6575*.

Yang, H.-W., Y. Zou, P. Shi, W. Lu, J. Lin, and S. Xu. (2019). "Aligning Cross-Lingual Entities with Multi-Aspect Information". 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)*. 4422–4432.

Yang, K. and J. Deng. (2020). "Strongly Incremental Constituency Parsing with Graph Neural Networks". *arXiv preprint arXiv:2010.14568*.

Yang, L., F. Wu, J. Gu, C. Wang, X. Cao, D. Jin, and Y. Guo. (2020). "Graph Attention Topic Modeling Network". In: *Proceedings of The Web Conference 2020. WWW '20*. New York, NY, USA: Association for Computing Machinery. 144–154. DOI: 10.1145/3366423.3380102. (Accessed on 12/25/2020).

Yang, L., Q. Ai, J. Guo, and W. B. Croft. (2016). "aNMM: Ranking short answer texts with attention-based neural matching model". In: *Proceedings of the 25th ACM international on conference on information and knowledge management*. 287–296.

Yang, Z., P. Qi, S. Zhang, Y. Bengio, W. Cohen, R. Salakhutdinov, and C. D. Manning. (2018a). "HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. 2369–2380.

Yang, Z., J. Zhao, B. Dhingra, K. He, W. W. Cohen, R. R. Salakhutdinov, and Y. LeCun. (2018b). "Glomo: Unsupervised learning of transferable relational graphs". In: *Advances in Neural Information Processing Systems*. 8950–8961.

Yao, L., C. Mao, and Y. Luo. (2019a). "Graph Convolutional Networks for Text Classification". In: *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*. AAAI Press. 7370–7377.

Yao, L., C. Mao, and Y. Luo. (2019b). "Graph convolutional networks for text classification". In: *Proc. AAAI Conf. Artif. Intell.* Vol. 33. 7370–7377.

Yao, S., T. Wang, and X. Wan. (2020). "Heterogeneous graph transformer for graph-to-sequence learning". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 7145–7154.

Yao, T., Y. Pan, Y. Li, and T. Mei. (2018). "Exploring visual relationship for image captioning". In: *Proceedings of the European conference on computer vision (ECCV)*. 684–699.

Yasunaga, M., H. Ren, A. Bosselut, P. Liang, and J. Leskovec. (2021). "QA-GNN: Reasoning with Language Models and Knowledge Graphs for Question Answering".

Yasunaga, M., R. Zhang, K. Meelu, A. Pareek, K. Srinivasan, and D. Radev. (2017). "Graph-based Neural Multi-Document Summarization". In: *Proceedings of the 21st Conference on Computational Natural Language Learning (CoNLL 2017)*. Vancouver, Canada: Association for Computational Linguistics. 452–462. DOI: 10.18653/v1/K17-1045.

Ye, R., X. Li, Y. Fang, H. Zang, and M. Wang. (2019). "A Vectorized Relational Graph Convolutional Network for Multi-Relational Network Alignment." In: *IJCAI*. 4135–4141.

Yih, W.-t., M.-W. Chang, X. He, and J. Gao. (2015). "Semantic Parsing via Staged Query Graph Generation: Question Answering with Knowledge Base". In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. 1321–1331.

Yih, W.-t., M. Richardson, C. Meek, M.-W. Chang, and J. Suh. (2016). "The value of semantic parse labeling for knowledge base question answering". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. 201–206.

Yin, P., G. Neubig, M. Allamanis, M. Brockschmidt, and A. L. Gaunt. (2018). "Learning to represent edits". *arXiv preprint arXiv:1810.13337*.

Yin, Y., F. Meng, J. Su, C. Zhou, Z. Yang, J. Zhou, and J. Luo. (2020). "A Novel Graph-based Multi-modal Fusion Encoder for Neural Machine Translation". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 3025–3035. DOI: 10.18653/v1/2020.acl-main.273.

Ying, R., J. You, C. Morris, X. Ren, W. L. Hamilton, and J. Leskovec. (2018). "Hierarchical graph representation learning with differentiable pooling". *arXiv preprint arXiv:1806.08804*.

Yu, T., R. Zhang, K. Yang, M. Yasunaga, D. Wang, Z. Li, J. Ma, I. Li, Q. Yao, S. Roman, Z. Zhang, and D. Radev. (2018). "Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 3911–3921. DOI: 10.18653/v1/D18-1425.

Yun, S., M. Jeong, R. Kim, J. Kang, and H. J. Kim. (2019). "Graph Transformer Networks". *Advances in Neural Information Processing Systems*. 32: 11983–11993.

Zellers, R., Y. Bisk, R. Schwartz, and Y. Choi. (2018). "Swag: A large-scale adversarial dataset for grounded commonsense inference". *arXiv preprint arXiv:1808.05326*.

Zeng, S., R. Xu, B. Chang, and L. Li. (2020). “Double Graph Based Reasoning for Document-level Relation Extraction”. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 1630–1640. DOI: 10.18653/v1/2020.emnlp-main.127.

Zhang, B., Y. Zhang, R. Wang, Z. Li, and M. Zhang. (2020a). “Syntax-aware opinion role labeling with dependency graph convolutional networks”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 3249–3258.

Zhang, C., Q. Li, and D. Song. (2019a). “Aspect-based Sentiment Classification with Aspect-specific Graph Convolutional Networks”. 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)*. Hong Kong, China: Association for Computational Linguistics. 4568–4578. DOI: 10.18653/v1/D19-1464. (Accessed on 12/26/2020).

Zhang, C., D. Song, C. Huang, A. Swami, and N. V. Chawla. (2019b). “Heterogeneous graph neural network”. In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 793–803.

Zhang, C., D. Song, C. Huang, A. Swami, and N. V. Chawla. (2019c). “Heterogeneous Graph Neural Network”. In: *KDD*.

Zhang, J., L. Wang, R. K.-W. Lee, Y. Bin, Y. Wang, J. Shao, and E.-P. Lim. (2020b). “Graph-to-Tree Learning for Solving Math Word Problems”. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 3928–3937. DOI: 10.18653/v1/2020.acl-main.362.

Zhang, M. and T. Qian. (2020). “Convolution over Hierarchical Syntactic and Lexical Graphs for Aspect Level Sentiment Analysis”. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 3540–3549. DOI: 10.18653/v1/2020.emnlp-main.286.

Zhang, N., S. Deng, J. Li, X. Chen, W. Zhang, and H. Chen. (2020c). "Summarizing Chinese Medical Answer with Graph Convolution Networks and Question-focused Dual Attention". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Online: Association for Computational Linguistics. 15–24. DOI: 10.18653/v1/2020.findings-emnlp.2.

Zhang, N., S. Deng, Z. Sun, G. Wang, X. Chen, W. Zhang, and H. Chen. (2019d). "Long-tail Relation Extraction via Knowledge Graph Embeddings and Graph Convolution Networks". In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics. 3016–3025. DOI: 10.18653/v1/N19-1306.

Zhang, S., X. Liu, J. Liu, J. Gao, K. Duh, and B. Van Durme. (2018a). "Record: Bridging the gap between human and machine commonsense reading comprehension". *arXiv preprint arXiv:1810.12885*.

Zhang, S., X. Ma, K. Duh, and B. Van Durme. (2019e). "AMR Parsing as Sequence-to-Graph Transduction". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 80–94.

Zhang, S., X. Ma, K. Duh, and B. Van Durme. (2019f). "AMR Parsing as Sequence-to-Graph Transduction". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 80–94. DOI: 10.18653/v1/P19-1009.

Zhang, X., J. Zhao, and Y. LeCun. (2015). "Character-level convolutional networks for text classification". *arXiv preprint arXiv:1509.01626*.

Zhang, Y., Z. Guo, Z. Teng, W. Lu, S. B. Cohen, Z. Liu, and L. Bing. (2020d). "Lightweight, Dynamic Graph Convolutional Networks for AMR-to-Text Generation". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 2162–2172. DOI: 10.18653/v1/2020.emnlp-main.169.

Zhang, Y., W. Chan, and N. Jaitly. (2017a). "Very deep convolutional networks for end-to-end speech recognition". In: *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. 4845–4849.

Zhang, Y., Q. Liu, and L. Song. (2018b). "Sentence-State LSTM for Text Representation". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 317–327.

Zhang, Y., X. Yu, Z. Cui, S. Wu, Z. Wen, and L. Wang. (2020e). "Every Document Owns Its Structure: Inductive Text Classification via Graph Neural Networks". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*. Ed. by D. Jurafsky, J. Chai, N. Schluter, and J. R. Tetreault. Association for Computational Linguistics. 334–339.

Zhang, Y., X. Yu, Z. Cui, S. Wu, Z. Wen, and L. Wang. (2020f). "Every document owns its structure: Inductive text classification via graph neural networks". *arXiv preprint arXiv:2004.13826*.

Zhang, Y., P. Qi, and C. D. Manning. (2018c). "Graph Convolution over Pruned Dependency Trees Improves Relation Extraction". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 2205–2215. DOI: 10.18653/v1/D18-1244.

Zhang, Y., V. Zhong, D. Chen, G. Angeli, and C. D. Manning. (2017b). "Position-aware attention and supervised data improve slot filling". In: *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 35–45.

Zhang, Y., H. Dai, Z. Kozareva, A. Smola, and L. Song. (2018d). "Variational reasoning for question answering with knowledge graph". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. No. 1.

Zhang, Z., F. Zhuang, H. Zhu, Z.-P. Shi, H. Xiong, and Q. He. (2020g). "Relational Graph Neural Network with Hierarchical Attention for Knowledge Graph Completion." In: *AAAI*. 9612–9619.

Zhao, J., X. Wang, C. Shi, B. Hu, G. Song, and Y. Ye. (2021). "Heterogeneous Graph Structure Learning for Graph Neural Networks". In: *Proceedings of the AAAI Conference on Artificial Intelligence*.

Zhao, L., W. Xu, and J. Guo. (2020a). "Improving Abstractive Dialogue Summarization with Graph Structures and Topic Words". In: *Proceedings of the 28th International Conference on Computational Linguistics*. Barcelona, Spain (Online): International Committee on Computational Linguistics. 437–449. URL: https://www.aclweb.org/anthology/2020.coling-main.39.

Zhao, Y., L. Chen, Z. Chen, R. Cao, S. Zhu, and K. Yu. (2020b). "Line Graph Enhanced AMR-to-Text Generation with Mix-Order Graph Attention Networks". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 732–741. DOI: 10.18653/v1/2020.acl-main.67.

Zhao, Y., L. Xiang, J. Zhu, J. Zhang, Y. Zhou, and C. Zong. (2020c). "Knowledge Graph Enhanced Neural Machine Translation via Multitask Learning on Sub-entity Granularity". In: *Proceedings of the 28th International Conference on Computational Linguistics*. Barcelona, Spain (Online): International Committee on Computational Linguistics. 4495–4505. URL: https://www.aclweb.org/anthology/2020.coling-main.397.

Zheng, B., H. Wen, Y. Liang, N. Duan, W. Che, D. Jiang, M. Zhou, and T. Liu. (2020). "Document Modeling with Graph Attention Networks for Multi-grained Machine Reading Comprehension". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, ACL 2020, Online, July 5-10, 2020*. Ed. by D. Jurafsky, J. Chai, N. Schluter, and J. R. Tetreault. Association for Computational Linguistics. 6708–6718.

Zheng, C. and P. Kordjamshidi. (2020). "SRLGRN: Semantic Role Labeling Graph Reasoning Network". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 8881–8891.

Zhong, V., C. Xiong, and R. Socher. (2017). "Seq2sql: Generating structured queries from natural language using reinforcement learning". *arXiv preprint arXiv:1709.00103*.

Zhou, D., X. Hu, and R. Wang. (2020a). "Neural Topic Modeling by Incorporating Document Relationship Graph". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Online: Association for Computational Linguistics. 3790–3796. DOI: 10.18653/v1/2020.emnlp-main.310. (Accessed on 12/26/2020).

Zhou, H., T. Young, M. Huang, H. Zhao, J. Xu, and X. Zhu. (2018a). "Commonsense knowledge aware conversation generation with graph attention." In: *IJCAI*. 4623–4629.

Zhou, Q., Y. Zhang, D. Ji, and H. Tang. (2020b). "AMR Parsing with Latent Structural Information". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. 4306–4319. DOI: 10.18653/v1/2020.acl-main.397.

Zhou, Q., N. Yang, F. Wei, S. Huang, M. Zhou, and T. Zhao. (2018b). "Neural document summarization by jointly learning to score and select sentences". *arXiv preprint arXiv:1807.02305*.

Zhu, H., Y. Lin, Z. Liu, J. Fu, T.-S. Chua, and M. Sun. (2019a). "Graph Neural Networks with Generated Parameters for Relation Extraction". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics. 1331–1339. DOI: 10.18653/v1/P19-1128.

Zhu, J., J. Li, M. Zhu, L. Qian, M. Zhang, and G. Zhou. (2019b). "Modeling Graph Structure in Transformer for Better AMR-to-Text Generation". 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)*. Hong Kong, China: Association for Computational Linguistics. 5459–5468. DOI: 10.18653/v1/D19-1548.

Zhu, J., J. Li, M. Zhu, L. Qian, M. Zhang, and G. Zhou. (2019c). "Modeling Graph Structure in Transformer for Better AMR-to-Text Generation". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. 5458–5467.

Zhu, Q., Z. Feng, and X. Li. (2018). "GraphBTM: Graph Enhanced Autoencoded Variational Inference for Biterm Topic Model". In: *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Brussels, Belgium: Association for Computational Linguistics. 4663–4672. DOI: 10.18653/v1/D18-1495.