Question Answering over Curated and Open Web Sources

Rishiraj Saha Roy  
Max Planck Institute for Informatics  
Saarbrücken, Germany  
rishiraj@mpi2.de

Avishek Anand  
L3S Research Center  
Hannover, Germany  
anand@l3s.de

ABSTRACT
The last few years have seen an explosion of research on the topic of automated question answering (QA), spanning the communities of information retrieval, natural language processing, and artificial intelligence. This tutorial would cover the highlights of this really active period of growth for QA to give the audience a grasp over the families of algorithms that are currently being used. We partition research contributions by the underlying source from where answers are retrieved: curated knowledge graphs, unstructured text, or hybrid corpora. We choose this dimension of partitioning as it is the most discriminative when it comes to algorithm design. Other key dimensions are covered within each sub-topic: like the complexity of questions addressed, and degrees of explainability and interactivity introduced in the systems. We would conclude the tutorial with the most promising emerging trends in the expanse of QA, that would help new entrants into this field make the best decisions to take the community forward. This tutorial was recently presented at SIGIR 2020.

KEYWORDS
Question answering, Open-domain QA, Text-QA, Passage ranking, Knowledge graphs, KG-QA, Table-QA

Perspectives. In IR, QA was traditionally treated as a special use case in search [30], to provide crisp and direct answers to certain classes of queries, as an alternative to ranked lists of documents that users would have to sift through. Such queries, with objective answers, are often referred to as factoid questions [7, 9] (a term whose definition has evolved over the years). Factoid QA became very popular with the emergence of large curated knowledge graphs (KGs) like YAGO, DBpedia, Freebase and Wikidata, powerful resources that enable such crisp question answering at scale. Question answering over knowledge graphs or equivalently, knowledge bases (KG-QA or KB-QA) became a field of its own, that is producing an increasing number of research contributions year over year [2, 3, 5, 11, 22, 26, 29, 32]. Effort has also been directed at answering questions over Web tables [17, 21], that can be considered canonicalizations of the challenges in QA over structured KGs.

In contrast, QA in NLP started with the AI goal of whether machines can comprehend simple passages [4, 6, 23, 34] so as to be able to answer questions posed from the contents of these passages. Over time, this machine reading comprehension (MRC) task became coupled with the retrieval pipeline, resulting in the so-called paradigm of open-domain QA [4, 10, 31] (a term that is overloaded with other senses as well [1, 12]). Nevertheless, this introduction of the retrieval pipeline led to a revival of text-QA, that had increasingly focused on non-factoid QA [8, 33] after the rise of structured KGs. This has also helped bridge the gap between text and KG-QA, with the latter family gradually incorporating supplementary textual sources to boost recall [24, 25, 27, 28]. Considering such heterogeneous sources may often be the right choice owing to the fact that KGs, while capturing an impressive amount of objective world knowledge, are inherently incomplete.

Background. Over several decades, the field of question answering (QA) grew steadily from early prototypes like BASEBALL [14], through IBM Watson [13] and all the way to present-day integration in virtually all personal assistants like Siri, Cortana, Alexa, and the Google Assistant. In the last few years though, research on QA has well and truly exploded: this has often resulted in top conferences regularly creating submission tracks and presentation sessions dedicated to this topic. This tutorial will try to highlight key contributions to automated QA systems in the last three to four years coming from the perspectives of information retrieval (IR) and natural language processing (NLP) [4–6, 10, 15, 19, 22, 23, 28, 29, 32].

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

Proceedings of the 2020 ACM SIGIR International Conference on the Theory of Information Retrieval (ICTIR ’20), September 14–17, 2020, Virtual Event, Norway. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3409256.3409809

1 OVERVIEW

Objectives. The goal of this tutorial is to give the audience a feel of commonalities in methods, challenges and opportunities in QA across different paradigms: this can have a significant effect on overcoming the severely fragmented view of the QA community.

Perspectives. In IR, QA was traditionally treated as a special use case in search [30], to provide crisp and direct answers to certain classes of queries, as an alternative to ranked lists of documents that users would have to sift through. Such queries, with objective answers, are often referred to as factoid questions [7, 9] (a term whose definition has evolved over the years). Factoid QA became very popular with the emergence of large curated knowledge graphs (KGs) like YAGO, DBpedia, Freebase and Wikidata, powerful resources that enable such crisp question answering at scale. Question answering over knowledge graphs or equivalently, knowledge bases (KG-QA or KB-QA) became a field of its own, that is producing an increasing number of research contributions year over year [2, 3, 5, 11, 22, 26, 29, 32]. Effort has also been directed at answering questions over Web tables [17, 21], that can be considered canonicalizations of the challenges in QA over structured KGs.

In contrast, QA in NLP started with the AI goal of whether machines can comprehend simple passages [4, 6, 23, 34] so as to be able to answer questions posed from the contents of these passages. Over time, this machine reading comprehension (MRC) task became coupled with the retrieval pipeline, resulting in the so-called paradigm of open-domain QA [4, 10, 31] (a term that is overloaded with other senses as well [1, 12]). Nevertheless, this introduction of the retrieval pipeline led to a revival of text-QA, that had increasingly focused on non-factoid QA [8, 33] after the rise of structured KGs. This has also helped bridge the gap between text and KG-QA, with the latter family gradually incorporating supplementary textual sources to boost recall [24, 25, 27, 28]. Considering such heterogeneous sources may often be the right choice owing to the fact that KGs, while capturing an impressive amount of objective world knowledge, are inherently incomplete.

In this tutorial, we refer to knowledge graphs and Web tables as the curated Web, and all unstructured text available online as the open Web.

Format and support. A detailed structure for our proposed half-day tutorial is available at https://www.avishekanand.com/talk/sigir20-tute/. It will last three hours plus breaks, and is aimed to meet a high-quality presentation within the chosen time period.
REFERENCES

[1] Abdalghani Abujabal, Rishiraj Saha Roy, Mohamed Yahya, and Gerhard Weikum. 2018. Never-ending learning for open-domain question answering over knowledge bases. In WWW.

[2] Nikita Bhutani, Xinyi Zheng, and HV Jagadish. 2019. Learning to Answer Complex Questions over Knowledge Bases with Query Composition. In CIKM.

[3] Soumen Chakrabarti. 2020. Interpretable Complex Question Answering. In NeurIPS.

[4] Charles L. Clarke and Egidio L. Terra. 2003. Passage retrieval vs. document retrieval for factoid question answering. In SIGIR.

[5] Mostafa Dehghani, Hosein Azarbonyad, Jaap Kamps, and Maarten de Rijke. 2019. Learning to transform, combine, and reason in open-domain question answering. In EMNLP-IJCNLP.

[6] Ahmed Elgohary, Chen Zhao, and Jordan Boyd-Graber. 2018. A dataset and baselines for sequential open-domain question answering. In EMNLP-IJCNLP.

[7] Charles L. A. Clarke and Egidio L. Terra. 2003. Passage retrieval vs. document retrieval for factoid question answering. In SIGIR.

[8] Nikita Bhutani, Xinyi Zheng, and HV Jagadish. 2019. Learning to Answer Complex Questions over Knowledge Bases with Query Composition. In CIKM.

[9] Silviu Cucerzan and Eugene Agichtein. 2005. Factoid Question Answering over Unstructured and Structured Web Content. In TREC.

[10] Christopher Clark and Matt Gardner. 2018. Simple and Effective Multi-Paragraph Reading Comprehension. In ACL.

[11] Jiwei Ding, Wei Hu, Qixin Xu, and Yuzhong Qu. 2019. Leveraging Frequent Query Answering over Knowledge Graphs Using Judicious Context Expansion. In CIKM.

[12] Mostafa Dehghani, Hosein Azarbonyad, Jaap Kamps, and Maarten de Rijke. 2019. Learning to transform, combine, and reason in open-domain question answering. In WSDM.

[13] Ahmed Elgohary, Chen Zhao, and Jordan Boyd-Graber. 2018. A dataset and baselines for sequential open-domain question answering. In EMNLP-IJCNLP.

[14] David Ferrucci, Eric Brown, Jennifer Chu-Carroll, James Fan, David Gondek, Aditya A. Kalyanpur, Adam Lally, J. William Murdock, Eric Nyberg, John Prager, Nico Schlaefer, and Chris Welty. 2010. Building Watson: An overview of the DeepQA project. AI magazine 31, 3 (2010).

[15] Bert F Green Jr, Adam Fish, Jason Weston, and Antoine Bordes. 2017. Search-based neural structured learning for sequential question answering. In ACL.

[16] Yangyang Guo, Zhiyong Cheng, Liqiang Nie, Yibing Liu, Yinglong Wang, and Mohan Kankanhalli. 2019. Quantifying and Alleviating the Language Prior Problem in Visual Question Answering. In SIGIR.

[17] Fei Li and HV Jagadish. 2014. Constructing an interactive natural language interface for relational databases. In VLDB.

[18] Xiaolu Lu, Soumajit Pramanik, Rishiraj Saha Roy, Abdalghani Abujabal, Yafang Wang, and Gerhard Weikum. 2019. Answering Complex Questions by Joining Multi-Document Evidence with Quasi Knowledge Graphs. In SIGIR.

[19] Anusri Pampari, Preethi Raghavan, Jennifer Liang, and Jian Peng. 2018. emrQA: A Large Corpus for Question Answering on Electronic Medical Records. In EMNLP.

[20] Anusri Pampari, Preethi Raghavan, Jennifer Liang, and Jian Peng. 2018. emrQA: A Large Corpus for Question Answering on Electronic Medical Records. In EMNLP.

[21] Panupong Pasupat and Percy Liang. 2015. Compositional Semantic Parsing over Semi-Structured Tables. In ACL.

[22] Yuanqi Qiu, Yuanzhao Wang, Xiaolong Jin, and Kun Zhang. 2020. Stepwise Reasoning for Multi-Relation Question Answering over Knowledge Graph with Weak Supervision. In WSDM.

[23] Pranav Rajpurkar, Jian Zhang, Konstantin Lopyrev, and Percy Liang. 2016. SQuAD: 100,000+ Questions for Machine Comprehension of Text. In EMNLP.

[24] Denis Savenkov and Eugene Agichtein. 2016. When a knowledge base is not enough: Question answering over knowledge bases with external text data. In SIGIR.

[25] Uma Sawant, Saurabh Garg, Soumen Chakrabarti, and Ganesh Ramakrishnan. 2019. Neural architecture for question answering using a knowledge graph and Web corpus. Information Retrieval Journal (2019).

[26] Tao Shen, Xiubo Geng, Qin Tao, Daya Guo, Duyu Tang, Nan Duan, Guodong Long, and Daxin Jiang. 2019. Multi-Task Learning for Conversational Question Answering over a Large-Scale Knowledge Base. In EMNLP-IJCNLP.

[27] Haitian Sun, Tania Bedrax-Weiss, and William Cohen. 2019. PullNet: Open Domain Question Answering with Iterative Retrieval on Knowledge Bases and Text. In EMNLP-IJCNLP.

[28] Haitian Sun, Bhuvan Dinha, Manzil Zaheer, Kathyitz Mazzaitis, Ruslan Salakhutdinov, and William Cohen. 2018. Open Domain Question Answering Using Early Fusion of Knowledge Bases and Text. In EMNLP.

[29] Svetlana Vakulenko, Javier David Fernandez Garcia, Axel Polleres, Maarten de Rijke, and Gerhard Weikum. 2019. Look before you Hop: Conversational Question Answering over Knowledge Graphs. In CIKM.

[30] Ellen M. Voorhees. 1999. The TREC-8 question answering track report. In TREC.

[31] Yingying Wang, Fei Li, Xiaolong Jin, and Kun Zhang. 2019. Document Gated Reader for Open-Domain Question Answering. In EMNLP-IJCNLP.

[32] Zhiyong Wu, Ben Kao, Tien-Hsuan Wu, Pengcheng Yin, and Qun Liu. 2020. PERQ: Predicting, Explaining, and Rectifying Failed Questions in KB-QA Systems. In SIGIR.

[33] Zhiyun Yang, Peng Qi, Saizheng Zhang, Yoshua Bengio, William Cohen, Ruslan Salakhutdinov, and Christopher D Manning. 2018. HotpotQA: A Dataset for Diverse, Explainable Multi-hop Question Answering. In EMNLP.