Question Answering and Generation from Structured and Unstructured Data

Yu Chen
Department of Computer Science
Rensselaer Polytechnic Institute
06/24/2020

Committee:
Prof. Mohammed J. Zaki (Chair), Prof. Bulent Yener
Prof. Alex Gittens, Prof. Qiang Ji, and Dr. Lingfei Wu

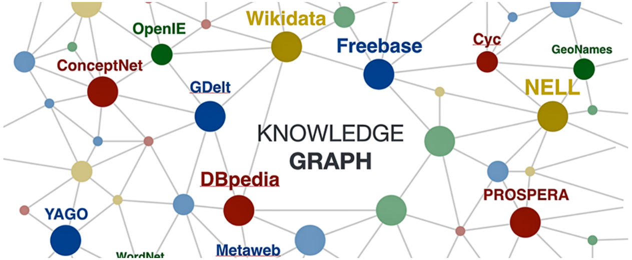
Rensselaer
why not change the world?®

Outline

- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

What is Question Answering?

Automated question answering (QA) is the process of finding answers to natural language questions using certain knowledge sources.



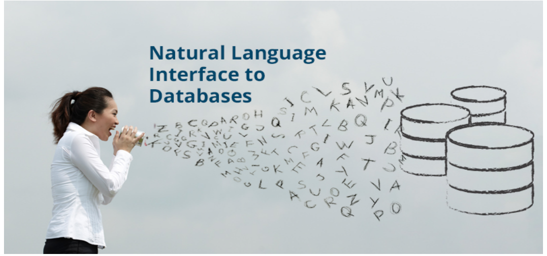
Ref: https://www.csee.umbc.edu/courses/graduate/691/fall19/07/

Ref: https://images.app.goo.gl/6pshxSDujQLTqnMK8

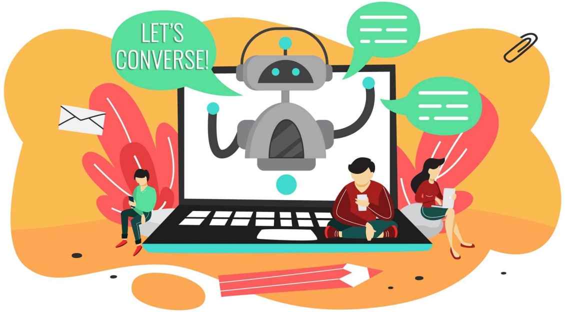
Ref: https://images.app.goo.gl/qvEnqksAACwqercBA

Question Answering Applications

- Natural language interfaces to databases
- Spoken dialog systems
- Beyond search engines



Ref: https://8kmiles.com/blog/natural-language-interface-databases/



https://images.app.goo.gl/ViagZT2oG1QERsR89

https://images.app.goo.gl/agDJmSVg5KaKc5o6A

Challenges of QA

Lexical gap

- Lexical gap between the question and context.
- Most previous methods ignore the subtle inter-relationships between the question and context.

Complex reasoning

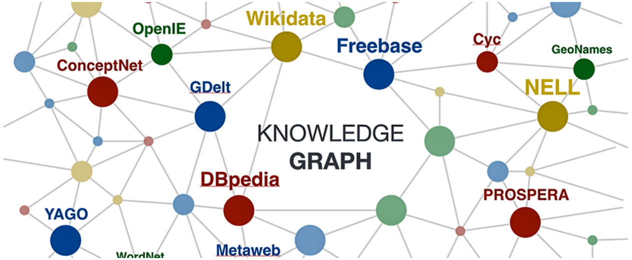
- Multi-hop reasoning.
- Diverse constraints.
- Many previous methods focus on single-hop QA without modeling various constraints.

Conversational QA

- Sequential questions.
- Most previous methods do not effectively capture conversation history.

What is Question Generation?

Natural question generation (QG) is the task of generating natural language questions from certain knowledge sources.



Ref: https://www.csee.umbc.edu/courses/graduate/691/fall19/07/

Ref: https://images.app.goo.gl/6pshxSDujQLTqnMK8

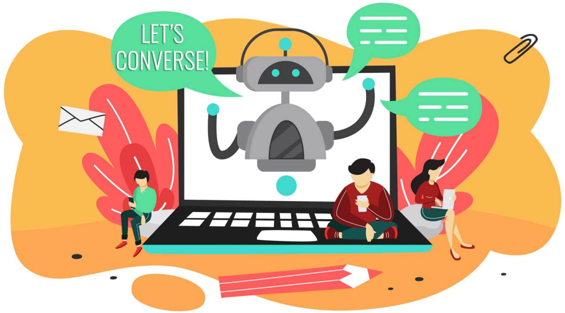
Ref: https://images.app.goo.gl/qvEnqksAACwqercBA

Question Generation Applications

- Improving the QA task by providing more training data
- Generating practice exercises for educational purposes
- Helping dialog systems

https://images.app.goo.gl/Nhb5toJWUwPnVT6Y7

https://images.app.goo.gl/FqaxidYkmBN7BMHm7



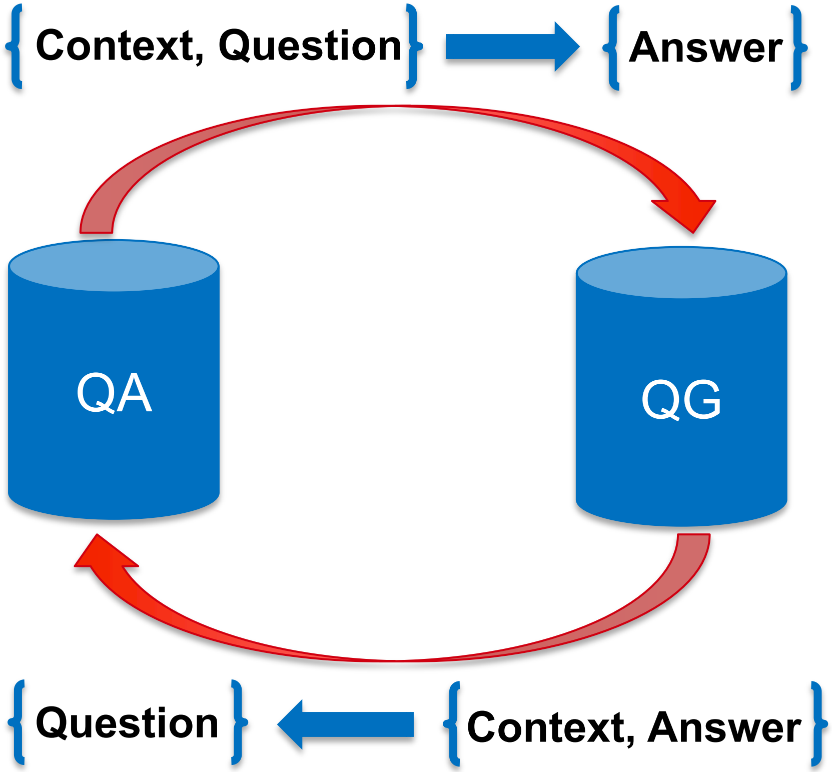
https://images.app.goo.gl/ViagZT2oG1QERsR89

Challenges of QG

Context modeling	Answer utilization	Model training
▪ Modeling long/large context. ▪ Modeling structure information in context. ▪ Most previous methods focus on short/small context and do not utilize rich structure info of context.	▪ Answer info for guiding the generation of relevant and meaningful questions. ▪ Most previous methods either do not consider or fail to effectively utilize answer info.	▪ Cross-entropy based sequence training has limitations. ▪ Most previous methods rely on cross-entropy loss or simple reinforcement learning (RL) loss for training.

QA & QG as Dual Tasks

- The input and the output of QA and QG are (almost) reverse.
- QA and QG can help improve each other.



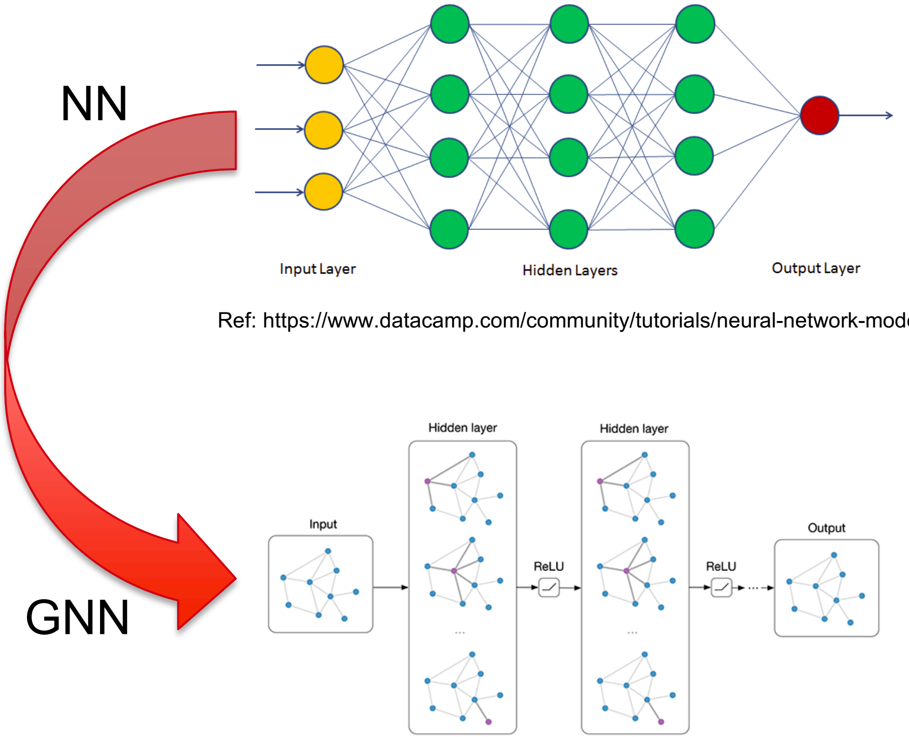
Outline

- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

GNN: Toward Geometric Deep Learning

- Graph Neural Networks (GNNs) generalize (structured) deep neural models to non-Euclidean domains such as graphs.

- GNNs have been widely applied to many tasks/domains:
  - Computer vision (CV)
  - Natural language processing (NLP)
  - Drug discovery
  - Social network analysis
  - Recommendation systems

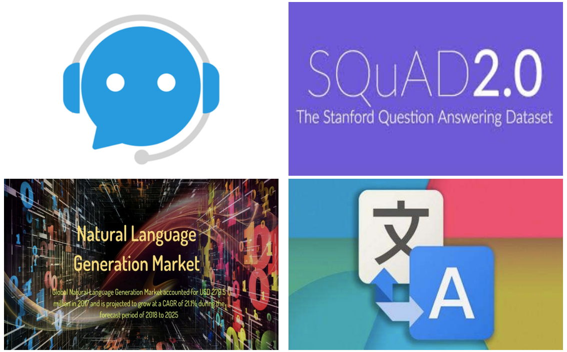


Ref: https://www.datacamp.com/community/tutorials/neural-network-models-r

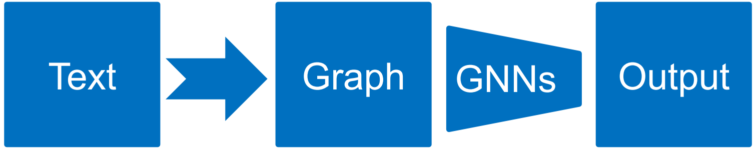
Ref: https://tkipf.github.io/graph-convolutional-networks/

GNN for NLP

- Applications
  - Dialog systems
  - Machine comprehension
  - X-to-text generation (e.g., AMR, SQL, etc.)
  - Machine translation

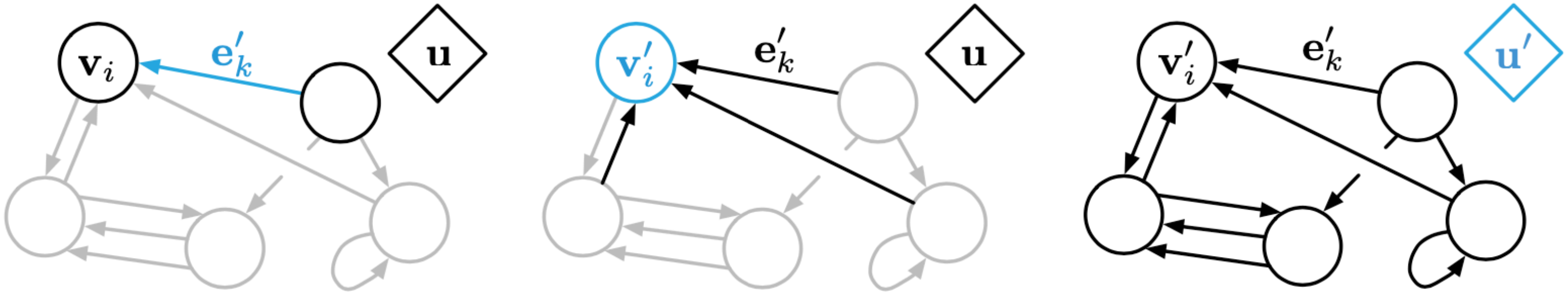


- Challenges
  - Only applicable to problems with graph-structured input.
  - Converting non-graph input to graph-structured input is crucial and non-trivial.



GNN: A Message Passing Perspective

- Computational steps within a GNN
  - (a) Edge update: passing node information to neighboring edges.
  - (b) Node update: for each node, aggregating information from neighboring edges.
  - (c) Global update: aggregating all node information.



(a) Edge update (b) Node update (c) Global update

Ref: Battaglia, Peter W., et al. "Relational inductive biases, deep learning, and graph networks." *arXiv preprint arXiv:1806.01261* (2018).

Outline

- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

Contributions: QA

QA

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

Lexical gap

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.

Lexical gap

- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.

Lexical gap

Complex reasoning

- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.

Lexical gap

Complex reasoning

- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

Lexical gap

Complex reasoning

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.

Lexical gap

Complex reasoning

- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Lexical gap

Complex reasoning

Conversation

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Lexical gap

Complex reasoning

Conversation

Contributions: QA

QA

- KBQA (NAACL 2019, ISWC 2019)
  - Modeling the two-way flow of interactions between the questions and the KB.
  - Multi-hop reasoning in a KB requiring no external resources and very few hand-crafted features.
  - Significantly outperforming previous IR-based methods while remaining competitive with handcrafted SP-based methods on a popular benchmark.
- Conversational machine comprehension (IJCAI 2020, ICML LRG 2019)
  - Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
  - Achieving superior performance compared to existing state-of-the-art methods on three public benchmarks.

Lexical gap

Complex reasoning

Conversation

Covered in candidacy exam

Contributions: QG

QG

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Context modeling

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Context modeling

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Context modeling

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Context modeling

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator. Context modeling
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator. Context modeling Model training
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator. Context modeling Model training
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator. Context modeling Model training
  - Deep Alignment Network for incorporating the answer information into the passage. Answer utilization
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph. — Context modeling
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator. — Context modeling; Model training
  - Deep Alignment Network for incorporating the answer information into the passage. — Answer utilization
  - Outperforming existing methods by a significant margin on a standard benchmark.

Contributions: QG

QG

- QG from KG (EMNLP 2020 under review)
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
  - Achieving new state-of-the-art scores on two benchmarks.
- QG from text (ICLR 2020, NeurIPS GRL 2019)
  - RL-based Graph2Seq model equipped with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.
  - Outperforming existing methods by a significant margin on a standard benchmark.

Context modeling

Model training

Context modeling

Answer utilization

Focus on QG

Research Achievements: QA

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Bidirectional Attentive Memory Networks for Question Answering over Knowledge Bases. In Proceedings of NAACL 2019.

Yu Chen, Lingfei Wu and Mohammed J. Zaki. GraphFlow: Exploiting Conversation Flow with Graph Neural Networks for Conversational Machine Comprehension. In Proceedings of IJCAI 2020.

Yu Chen, Lingfei Wu and Mohammed J. Zaki. GraphFlow: Exploiting Conversation Flow with Graph Neural Networks for Conversational Machine Comprehension. In ICML LRG 2019.

Steven Haussmann, Oshani Seneviratne, **Yu Chen**, Yarden Ne'eman, James Codella, Ching-Hua Chen, Deborah L. McGuinness and Mohammed J. Zaki. FoodKG: A Semantics-Driven Knowledge Graph for Food Recommendation. In Proceedings of ISWC 2019.

Steven Haussmann, **Yu Chen**, Oshani Seneviratne, Nidhi Rastogi, James Codella, Ching-Hua Chen, Deborah McGuinness, Mohammed J. Zaki. FoodKG Enabled Q&A Application. In Proceedings of ISWC 2019.

Yu Chen, Ananya Subburathinam, Ching-Hua Chen and Mohammed J. Zaki. Personalized Question Answering over a Large-scale Food Knowledge Graph. Submitted to RecSys 2020.

Research Achievements: QG

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Reinforcement Learning Based Graph-to-Sequence Model for Natural Question Generation. In Proceedings of ICLR 2020.

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Natural Question Generation with Reinforcement Learning Based Graph-to-Sequence Model. In NeurIPS GRL 2019.

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Toward Subgraph Guided Knowledge Graph Question Generation with Graph Neural Networks. Submitted to EMNLP 2020.

Research Achievements: Others

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Deep Iterative and Adaptive Learning for Graph Neural Networks. In The First International Workshop on Deep Learning on Graphs: Methodologies and Applications (with AAAI). February 2020. ***Best Student Paper Award***

Yu Chen and Mohammed J. Zaki. KATE: K-competitive Autoencoder for Text. In Proceedings of KDD 2017.

Yu Chen, Rhaad M. Rabbani, Aparna Gupta and Mohammed J. Zaki. Comparative Text Analytics via Topic Modeling in Banking. In Proceedings of IEEE SSCI 2017.

Yu Chen, Lingfei Wu and Mohammed J. Zaki. Iterative Deep Graph Learning for Graph Neural Networks: Better and Robust Node Embeddings. Submitted to NeurIPS 2020.

Outline

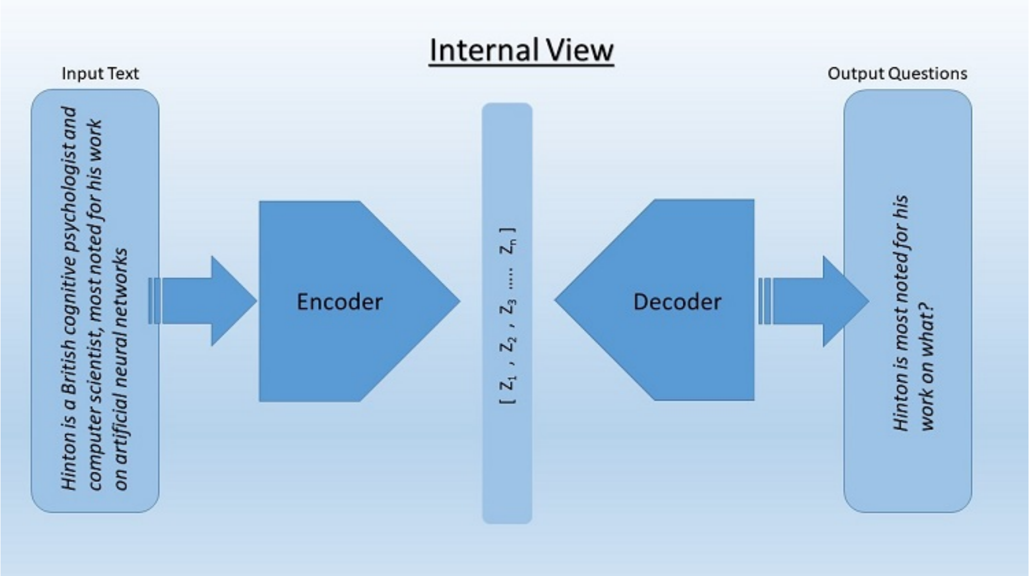
- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

Part I

Reinforcement Learning Based Graph-to-Sequence Model for Question Generation from Text

Yu Chen, Lingfei Wu, and Mohammed J. Zaki. "Reinforcement learning based graph-to-sequence model for natural question generation." *ICLR 2020.*

Problem Formulation



Ref: https://images.app.goo.gl/TJEqwFS8nBNW8n8SA

- Input:
  - A text passage $X^p = \{x_1^p, x_2^p, ..., x_N^p\}$
  - A target answer $X^a = \{x_1^a, x_2^a, ..., x_L^a\}$
- Output:
  - A natural language question

$$\hat{Y} = \{y_1, y_2, ..., y_T\}$$

which maximizes the conditional likelihood

$$\hat{Y} = \arg\max_Y P(Y|X^p, X^a)$$

Motivation

Context modeling

- Previous works: ignoring the rich structure information hidden in text.
- Our solution: applying a GNN-based encoder to capture rich structure information.

Answer utilization

- Previous works: Failing to fully exploit the answer information.
- Our solution: proposing a deep alignment network for attention-based soft alignment between passage and answer.

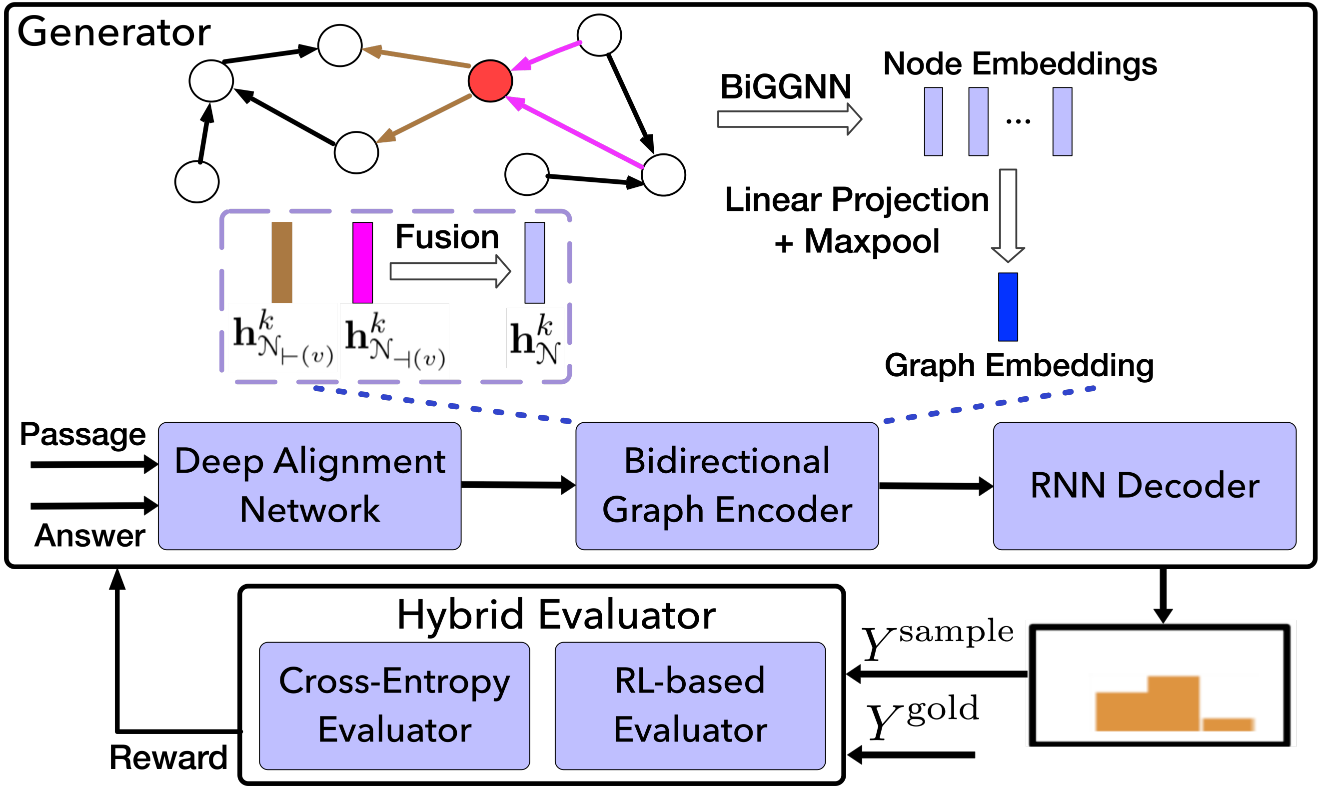
Model training

- Previous works: Solely relying on cross-entropy loss or simple RL loss for training.
- Our solution: designing a hybrid loss combining both cross-entropy loss and RL loss.

Contributions

- We propose a novel RL-based Graph2Seq model for question generation from text. To the best of our knowledge, we are the first to introduce the Graph2Seq architecture for QG.
- We design a novel deep alignment network to effectively utilize the answer information.
- We present a mixed loss function combining both cross-entropy loss and RL loss.
- We explore both static and dynamic ways of constructing graphs from text and are the first to systematically investigate their performance impacts on a GNN encoder.
- The proposed model outperforms existing methods by a significant margin on the standard SQuAD benchmark for QG.

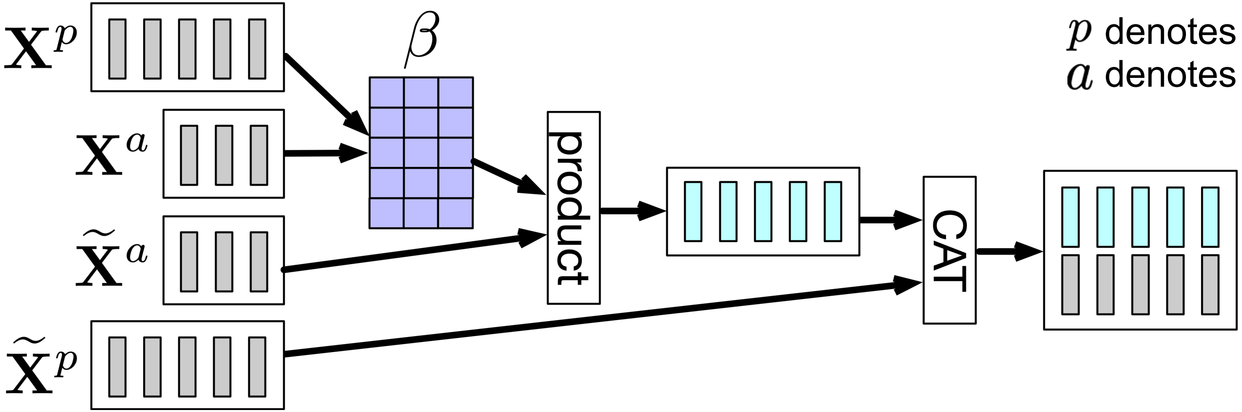
Overall Model Architecture



Deep Answer Alignment: Overview

- Previous methods neglect potential semantic relationships between passages and answers when utilizing answer information.
- We explicitly model the global interactions among passages and answers in the embedding space.
  - A deep alignment network for incorporating the answer information into passages with multiple granularity levels.
  - We perform attention-based soft-alignment at both the word level and the contextual level.

Deep Answer Alignment: Formulation



Key-value representations of passage and answer

p denotes passage
a denotes answer

Step 3: compute final passage embeddings

Step 2: compute aligned answer embeddings

$$\widetilde{\mathbf{H}}^p = \text{Align}(\mathbf{X}^p, \mathbf{X}^a, \widetilde{\mathbf{X}}^p, \widetilde{\mathbf{X}}^a) = \text{CAT}(\widetilde{\mathbf{X}}^p; \mathbf{H}^p) = \text{CAT}(\widetilde{\mathbf{X}}^p; \widetilde{\mathbf{X}}^a \boldsymbol{\beta}^T)$$

Step1: compute passage-answer attention matrix

$$\boldsymbol{\beta} \propto \exp\Big(\text{ReLU}(\mathbf{W}\mathbf{X}^p)^T \text{ReLU}(\mathbf{W}\mathbf{X}^a)\Big)$$

Deep Answer Alignment: Word Level

- On the passage side
  - We perform deep answer alignment between passage and answer based on their word embeddings to obtain the passage embeddings $\widetilde{\mathbf{H}}^p$.
  - A BiLSTM is applied to $\widetilde{\mathbf{H}}^p$ to obtain contextualized passage embeddings.
- On the answer side
  - A BiLSTM is applied to the answer word embedding sequence to obtain the contextualized answer embeddings.

Deep Answer Alignment: Contextual Level

- On the passage side
  - We perform deep answer alignment between passage and answer based on their contextualized embeddings.
  - A BiLSTM is applied to the above obtained passage embeddings to compute the final passage embeddings $\mathbf{X}$.

Generator: Passage Graph Construction

- Syntax-based static graph construction
  - A directed and unweighted passage graph based on dependency parsing.

- Semantics-aware dynamic graph construction
  - A directed and weighted graph modeling semantic relationships among passage words.

Attention matrix → $$\mathbf{A} = \mathrm{ReLU}(\mathbf{U}\widetilde{\mathbf{H}}^p)^T \, \mathrm{ReLU}(\mathbf{U}\widetilde{\mathbf{H}}^p)$$

Sparse attention matrix → $$\bar{\mathbf{A}} = \mathrm{kNN}(\mathbf{A})$$

Normalized attention matrix → $$\mathbf{A}^{\dashv}, \mathbf{A}^{\vdash} = \mathrm{softmax}(\{\bar{\mathbf{A}}, \bar{\mathbf{A}}^T\})$$

Passage word as node!

$\widetilde{\mathbf{H}}^p$ is the passage representation

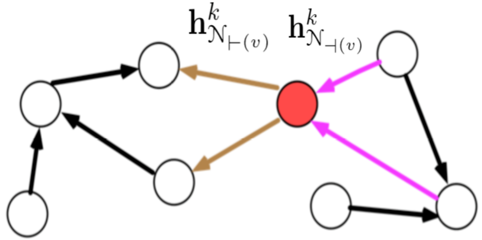
Node aggregation for the syntax-based static graph

$$\mathbf{h}^k_{\mathcal{N}_{\dashv(v)}} = \text{MEAN}(\{\mathbf{h}^{k-1}_v\} \cup \{\mathbf{h}^{k-1}_u, \forall u \in \mathcal{N}_{\dashv(v)}\})$$
$$\mathbf{h}^k_{\mathcal{N}_{\vdash(v)}} = \text{MEAN}(\{\mathbf{h}^{k-1}_v\} \cup \{\mathbf{h}^{k-1}_u, \forall u \in \mathcal{N}_{\vdash(v)}\})$$

Node embeddings are initialized to the passage embeddings $\mathbf{X}$ returned by Deep Alignment Network.

Node aggregation for the semantics-based static graph

$$\mathbf{h}^k_{\mathcal{N}_{\dashv(v)}} = \sum_{\forall u \in \mathcal{N}_{\dashv(v)}} \mathbf{a}^{\dashv}_{v,u} \mathbf{h}^{k-1}_u, \quad \mathbf{h}^k_{\mathcal{N}_{\vdash(v)}} = \sum_{\forall u \in \mathcal{N}_{\vdash(v)}} \mathbf{a}^{\vdash}_{v,u} \mathbf{h}^{k-1}_u$$



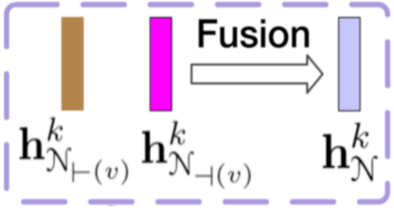
Generator: Bidirectional Gated GNN Encoder (Cont'd)

Fuse the aggregated node embeddings from both directions at each GNN hop

$$\mathbf{h}^k_{\mathcal{N}_{(v)}} = \mathrm{Fuse}(\mathbf{h}^k_{\mathcal{N}_{\dashv(v)}}, \mathbf{h}^k_{\mathcal{N}_{\vdash(v)}})$$

$$\mathrm{Fuse}(\mathbf{a}, \mathbf{b}) = \mathbf{z} \odot \mathbf{a} + (1 - \mathbf{z}) \odot \mathbf{b}$$

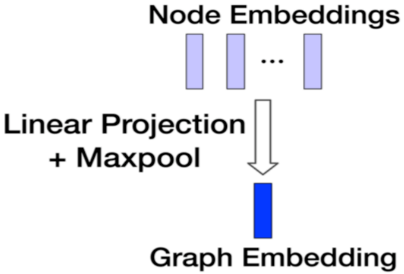
$$\mathbf{z} = \sigma(\mathbf{W}_z[\mathbf{a}; \mathbf{b}; \mathbf{a} \odot \mathbf{b}; \mathbf{a} - \mathbf{b}] + \mathbf{b}_z)$$



Update the node embeddings using fused information

$$\mathbf{h}^k_v = \mathrm{GRU}(\mathbf{h}^{k-1}_v, \mathbf{h}^k_{\mathcal{N}_{(v)}})$$

where GRU is a Gated Recurrent Unit (Cho et al., 2014).



After n hops of GNN computation, we obtain the final node/graph embeddings.

Cho, Kyunghyun, et al. "Learning phrase representations using RNN encoder-decoder for statistical machine translation." *arXiv preprint arXiv:1406.1078* (2014).

Generator: RNN Decoder

- We adopt a state-of-the-art attention-based LSTM decoder with copy and coverage mechanisms (See et al., 2017).
- Initial hidden states are based on graph embeddings.
- Node embeddings can be accessed via attention mechanism as a memory bank.

See, Abigail, Peter J. Liu, and Christopher D. Manning. "Get to the point: Summarization with pointer-generator networks." *arXiv preprint arXiv:1704.04368* (2017).

Hybrid Evaluator

- Cross-entropy based training has limitations, e.g., exposure bias.
- A mixed loss combining both cross-entropy loss and RL loss
  - Ensure the generation of syntactically and semantically valid text

$$\mathcal{L}_{lm} = \sum_t -\log P(y_t^* | X, y_{<t}^*) + \lambda \, \text{covloss}_t \qquad \mathcal{L}_{rl} = (r(\hat{Y}) - r(Y^s)) \sum_t \log P(y_t^s | X, y_{<t}^s)$$

$$\mathcal{L} = \gamma \mathcal{L}_{rl} + (1 - \gamma) \mathcal{L}_{lm}$$

greedy search

multinomial sampling

self-critical sequence training (SCST) RL algorithm.

- Two-stage training strategy:
  - Train the model with cross-entropy loss
  - Finetune the model by optimizing the mixed objective function

Experimental Setup: Data

- SQuAD:
  - Popular benchmark for the task of Machine Reading Comprehension.
- Our QG benchmarks:
  - SQuAD split 1: 75,500/17,934/11,805 (train/development/test) examples
  - SQuAD split 2: 86,635/8,965/8,964 examples

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under **gravity**. The main forms of precipitation include drizzle, rain, sleet, snow, **graupel** and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called “showers”.

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

Sample question-answer pairs from SQuAD 1.0. Ref: https://arxiv.org/abs/1606.05250.

Experimental Setup: Evaluation Metrics

- Automatic evaluation
  - BLEU-4
  - METEOR
  - ROUGE-L
  - Q-BLEU1
- Human evaluation
  - Syntactically correct
  - Semantically correct
  - Relevant

Experimental Results: Automatic Evaluation

Methods	Split-1				Split-2			
	BLEU-4	METEOR	ROUGE-L	Q-BLEU1	BLEU-4	METEOR	ROUGE-L	Q-BLEU1
Transformer	2.56	8.98	26.01	16.70	3.09	9.68	28.86	20.10
SeqCopyNet	–	–	–	–	13.02	–	44.00	–
NQG++	–	–	–	–	13.29	–	–	–
MPQG+R*	14.39	18.99	42.46	52.00	14.71	18.93	42.60	50.30
AFPQA	–	–	–	–	15.64	–	–	–
s2sa-at-mp-gsa	15.32	19.29	43.91	–	15.82	19.67	44.24	–
ASs2s	16.20	19.92	43.96	–	16.17	–	–	–
CGC-QG	–	–	–	–	17.55	21.24	44.53	–
$G2S_{dyn}$+BERT+RL	17.55	21.42	45.59	55.40	18.06	21.53	45.91	55.00
$G2S_{sta}$+BERT+RL	**17.94**	**21.76**	**46.02**	**55.60**	**18.30**	**21.70**	**45.98**	**55.20**

Table 1: Automatic evaluation results on the SQuAD test set. (higher scores indicate better results).

Experimental Results: Human Evaluation

Methods	Syntactically correct	Semantically correct	Relevant
MPQG+R*	4.34 (0.15)	4.01 (0.23)	3.21 (0.31)
G2S$_{sta}$+BERT+RL	4.41 (0.09)	4.31 (0.12)	3.79 (0.45)
Ground-truth	**4.74 (0.14)**	**4.74 (0.19)**	**4.25 (0.38)**

Table 2: Human evaluation results (± standard deviation) on the SQuAD split-2 test set. (higher scores indicate better results).

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
G2S_{dyn}+BERT+RL	18.06	G2S_{dyn} w/o feat	16.51
G2S_{sta}+BERT+RL	18.30	G2S_{sta} w/o feat	16.65
G2S_{sta}+BERT-fixed+RL	18.20	G2S_{dyn} w/o DAN	12.58
G2S_{dyn}+BERT	17.56	G2S_{sta} w/o DAN	12.62
G2S_{sta}+BERT	18.02	G2S_{sta} w/ DAN-word only	15.92
G2S_{sta}+BERT-fixed	17.86	G2S_{sta} w/ DAN-contextual only	16.07
G2S_{dyn}+RL	17.18	G2S_{sta} w/ GGNN-forward	16.53
G2S_{sta}+RL	17.49	G2S_{sta} w/ GGNN-backward	16.75
G2S_{dyn}	16.81	G2S_{sta} w/o BiGGNN, w/ Seq2Seq	16.14
G2S_{sta}	16.96	G2S_{sta} w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
G2S_{dyn}+BERT+RL	18.06	G2S_{dyn} w/o feat	16.51
G2S_{sta}+BERT+RL	18.30	G2S_{sta} w/o feat	16.65
G2S_{sta}+BERT-fixed+RL	18.20	G2S_{dyn} w/o DAN	12.58
G2S_{dyn}+BERT	17.56	G2S_{sta} w/o DAN	12.62
G2S_{sta}+BERT	18.02	G2S_{sta} w/ DAN-word only	15.92
G2S_{sta}+BERT-fixed	17.86	G2S_{sta} w/ DAN-contextual only	16.07
G2S_{dyn}+RL	17.18	G2S_{sta} w/ GGNN-forward	16.53
G2S_{sta}+RL	17.49	G2S_{sta} w/ GGNN-backward	16.75
G2S_{dyn}	16.81	G2S_{sta} w/o BiGGNN, w/ Seq2Seq	16.14
G2S_{sta}	16.96	G2S_{sta} w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
G2S_{dyn}+BERT+RL	18.06	G2S_{dyn} w/o feat	16.51
G2S_{sta}+BERT+RL	18.30	G2S_{sta} w/o feat	16.65
G2S_{sta}+BERT-fixed+RL	18.20	G2S_{dyn} w/o DAN	12.58
G2S_{dyn}+BERT	17.56	G2S_{sta} w/o DAN	12.62
G2S_{sta}+BERT	18.02	G2S_{sta} w/ DAN-word only	15.92
G2S_{sta}+BERT-fixed	17.86	G2S_{sta} w/ DAN-contextual only	16.07
G2S_{dyn}+RL	17.18	G2S_{sta} w/ GGNN-forward	16.53
G2S_{sta}+RL	17.49	G2S_{sta} w/ GGNN-backward	16.75
G2S_{dyn}	16.81	G2S_{sta} w/o BiGGNN, w/ Seq2Seq	16.14
G2S_{sta}	16.96	G2S_{sta} w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
G2S_{dyn}+BERT+RL	18.06	G2S_{dyn} w/o feat	16.51
G2S_{sta}+BERT+RL	18.30	G2S_{sta} w/o feat	16.65
G2S_{sta}+BERT-fixed+RL	18.20	G2S_{dyn} w/o DAN	12.58
G2S_{dyn}+BERT	17.56	G2S_{sta} w/o DAN	12.62
G2S_{sta}+BERT	18.02	G2S_{sta} w/ DAN-word only	15.92
G2S_{sta}+BERT-fixed	17.86	G2S_{sta} w/ DAN-contextual only	16.07
G2S_{dyn}+RL	17.18	G2S_{sta} w/ GGNN-forward	16.53
G2S_{sta}+RL	17.49	G2S_{sta} w/ GGNN-backward	16.75
G2S_{dyn}	16.81	G2S_{sta} w/o BiGGNN, w/ Seq2Seq	16.14
G2S_{sta}	16.96	G2S_{sta} w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Ablation Study

Methods	BLEU-4	Methods	BLEU-4
$G2S_{dyn}$+BERT+RL	18.06	$G2S_{dyn}$ w/o feat	16.51
$G2S_{sta}$+BERT+RL	18.30	$G2S_{sta}$ w/o feat	16.65
$G2S_{sta}$+BERT-fixed+RL	18.20	$G2S_{dyn}$ w/o DAN	12.58
$G2S_{dyn}$+BERT	17.56	$G2S_{sta}$ w/o DAN	12.62
$G2S_{sta}$+BERT	18.02	$G2S_{sta}$ w/ DAN-word only	15.92
$G2S_{sta}$+BERT-fixed	17.86	$G2S_{sta}$ w/ DAN-contextual only	16.07
$G2S_{dyn}$+RL	17.18	$G2S_{sta}$ w/ GGNN-forward	16.53
$G2S_{sta}$+RL	17.49	$G2S_{sta}$ w/ GGNN-backward	16.75
$G2S_{dyn}$	16.81	$G2S_{sta}$ w/o BiGGNN, w/ Seq2Seq	16.14
$G2S_{sta}$	16.96	$G2S_{sta}$ w/o BiGGNN, w/ GCN	14.47

Table 3: Ablation study on the SQuAD split-2 test set.

Experimental Results: Case Study

Passage:	for the successful execution of a project , effective planning is essential .
Gold:	what is essential for the successful execution of a project ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq):	what type of planning is essential for the project ?
$G2S_{sta}$ w/o DAN.:	what type of planning is essential for the successful execution of a project ?
$G2S_{sta}$:	what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT:	what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT+RL:	what is essential for the successful execution of a project ?
$G2S_{dyn}$+BERT+RL:	what is essential for the successful execution of a project ?
Passage:	the church operates three hundred sixty schools and institutions overseas .
Gold:	how many schools and institutions does the church operate overseas ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq):	how many schools does the church have ?
$G2S_{sta}$ w/o DAN.:	how many schools does the church have ?
$G2S_{sta}$:	how many schools and institutions does the church have ?
$G2S_{sta}$+BERT:	how many schools and institutions does the church have ?
$G2S_{sta}$+BERT+RL:	how many schools and institutions does the church operate ?
$G2S_{dyn}$+BERT+RL:	how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage: for the successful execution of a project , effective planning is essential .
Gold: what is essential for the successful execution of a project ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq): what type of planning is essential for the project ?
$G2S_{sta}$ w/o DAN.: what type of planning is essential for the successful execution of a project ?
$G2S_{sta}$: what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT: what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT+RL: what is essential for the successful execution of a project ?
$G2S_{dyn}$+BERT+RL: what is essential for the successful execution of a project ?

Passage: the church operates three hundred sixty schools and institutions overseas .
Gold: how many schools and institutions does the church operate overseas ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq): how many schools does the church have ?
$G2S_{sta}$ w/o DAN.: how many schools does the church have ?
$G2S_{sta}$: how many schools and institutions does the church have ?
$G2S_{sta}$+BERT: how many schools and institutions does the church have ?
$G2S_{sta}$+BERT+RL: how many schools and institutions does the church operate ?
$G2S_{dyn}$+BERT+RL: how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage: for the successful execution of a project , effective planning is essential .	
Gold: what is essential for the successful execution of a project ?	
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): what type of planning is essential for the project ?	
G2S$_{sta}$ w/o DAN.: what type of planning is essential for the successful execution of a project ?	
G2S$_{sta}$: what is essential for the successful execution of a project ?	
G2S$_{sta}$+BERT: what is essential for the successful execution of a project ?	
G2S$_{sta}$+BERT+RL: what is essential for the successful execution of a project ?	
G2S$_{dyn}$+BERT+RL: what is essential for the successful execution of a project ?	
Passage: the church operates three hundred sixty schools and institutions overseas .	
Gold: how many schools and institutions does the church operate overseas ?	
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): how many schools does the church have ?	
G2S$_{sta}$ w/o DAN.: how many schools does the church have ?	
G2S$_{sta}$: how many schools and institutions does the church have ?	
G2S$_{sta}$+BERT: how many schools and institutions does the church have ?	
G2S$_{sta}$+BERT+RL: how many schools and institutions does the church operate ?	
G2S$_{dyn}$+BERT+RL: how many schools does the church operate ?	

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage: for the successful execution of a project , effective planning is essential .
Gold: what is essential for the successful execution of a project ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq): what type of planning is essential for the project ?
$G2S_{sta}$ w/o DAN.: what type of planning is essential for the successful execution of a project ?
$G2S_{sta}$: what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT: what is essential for the successful execution of a project ?
$G2S_{sta}$+BERT+RL: what is essential for the successful execution of a project ?
$G2S_{dyn}$+BERT+RL: what is essential for the successful execution of a project ?

Passage: the church operates three hundred sixty schools and institutions overseas .
Gold: how many schools and institutions does the church operate overseas ?
$G2S_{sta}$ w/o BiGGNN (Seq2Seq): how many schools does the church have ?
$G2S_{sta}$ w/o DAN.: how many schools does the church have ?
$G2S_{sta}$: how many schools and institutions does the church have ?
$G2S_{sta}$+BERT: how many schools and institutions does the church have ?
$G2S_{sta}$+BERT+RL: how many schools and institutions does the church operate ?
$G2S_{dyn}$+BERT+RL: how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage:	for the successful execution of a project , effective planning is essential .
Gold:	what is essential for the successful execution of a project ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq):	what type of planning is essential for the project ?
G2S$_{sta}$ w/o DAN.:	what type of planning is essential for the successful execution of a project ?
G2S$_{sta}$:	what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT:	what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT+RL:	what is essential for the successful execution of a project ?
G2S$_{dyn}$+BERT+RL:	what is essential for the successful execution of a project ?
Passage:	the church operates three hundred sixty schools and institutions overseas .
Gold:	how many schools and institutions does the church operate overseas ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq):	how many schools does the church have ?
G2S$_{sta}$ w/o DAN.:	how many schools does the church have ?
G2S$_{sta}$:	how many schools and institutions does the church have ?
G2S$_{sta}$+BERT:	how many schools and institutions does the church have ?
G2S$_{sta}$+BERT+RL:	how many schools and institutions does the church operate ?
G2S$_{dyn}$+BERT+RL:	how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage: for the successful execution of a project , effective planning is essential .
Gold: what is essential for the successful execution of a project ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): what type of planning is essential for the project ?
G2S$_{sta}$ w/o DAN.: what type of planning is essential for the successful execution of a project ?
G2S$_{sta}$: what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT: what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT+RL: what is essential for the successful execution of a project ?
G2S$_{dyn}$+BERT+RL: what is essential for the successful execution of a project ?

Passage: the church operates three hundred sixty schools and institutions overseas .
Gold: how many schools and institutions does the church operate overseas ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): how many schools does the church have ?
G2S$_{sta}$ w/o DAN.: how many schools does the church have ?
G2S$_{sta}$: how many schools and institutions does the church have ?
G2S$_{sta}$+BERT: how many schools and institutions does the church have ?
G2S$_{sta}$+BERT+RL: how many schools and institutions does the church operate ?
G2S$_{dyn}$+BERT+RL: how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Experimental Results: Case Study

Passage: for the successful execution of a project , <u>effective planning</u> is essential .
Gold: what is essential for the successful execution of a project ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): what type of planning is essential for the project ?
G2S$_{sta}$ w/o DAN.: what type of planning is essential for the successful execution of a project ?
G2S$_{sta}$: what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT: what is essential for the successful execution of a project ?
G2S$_{sta}$+BERT+RL: what is essential for the successful execution of a project ?
G2S$_{dyn}$+BERT+RL: what is essential for the successful execution of a project ?

Passage: the church operates <u>three hundred sixty schools and institutions overseas</u> .
Gold: how many schools and institutions does the church operate overseas ?
G2S$_{sta}$ w/o BiGGNN (Seq2Seq): how many schools does the church have ?
G2S$_{sta}$ w/o DAN.: how many schools does the church have ?
G2S$_{sta}$: how many schools and institutions does the church have ?
G2S$_{sta}$+BERT: how many schools and institutions does the church have ?
G2S$_{sta}$+BERT+RL: how many schools and institutions does the church operate ?
G2S$_{dyn}$+BERT+RL: how many schools does the church operate ?

Table 4: Generated questions on SQuAD split-2 test set. Target answers are underlined.

Outline

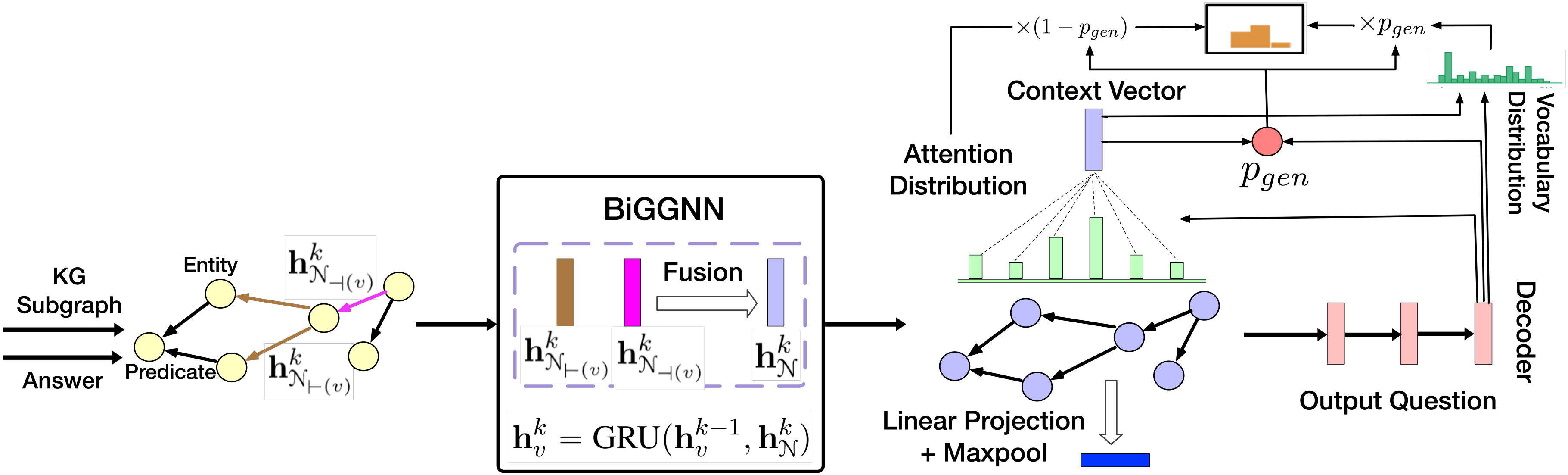
- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

Part II

Toward Subgraph Guided Knowledge Graph Question Generation with Graph Neural Networks

Yu Chen, Lingfei Wu, and Mohammed J. Zaki. "Toward Subgraph Guided Knowledge Graph Question Generation with Graph Neural Networks." Submitted to *EMNLP 2020.*

Graph2Seq for QG from KG: Overall Architecture

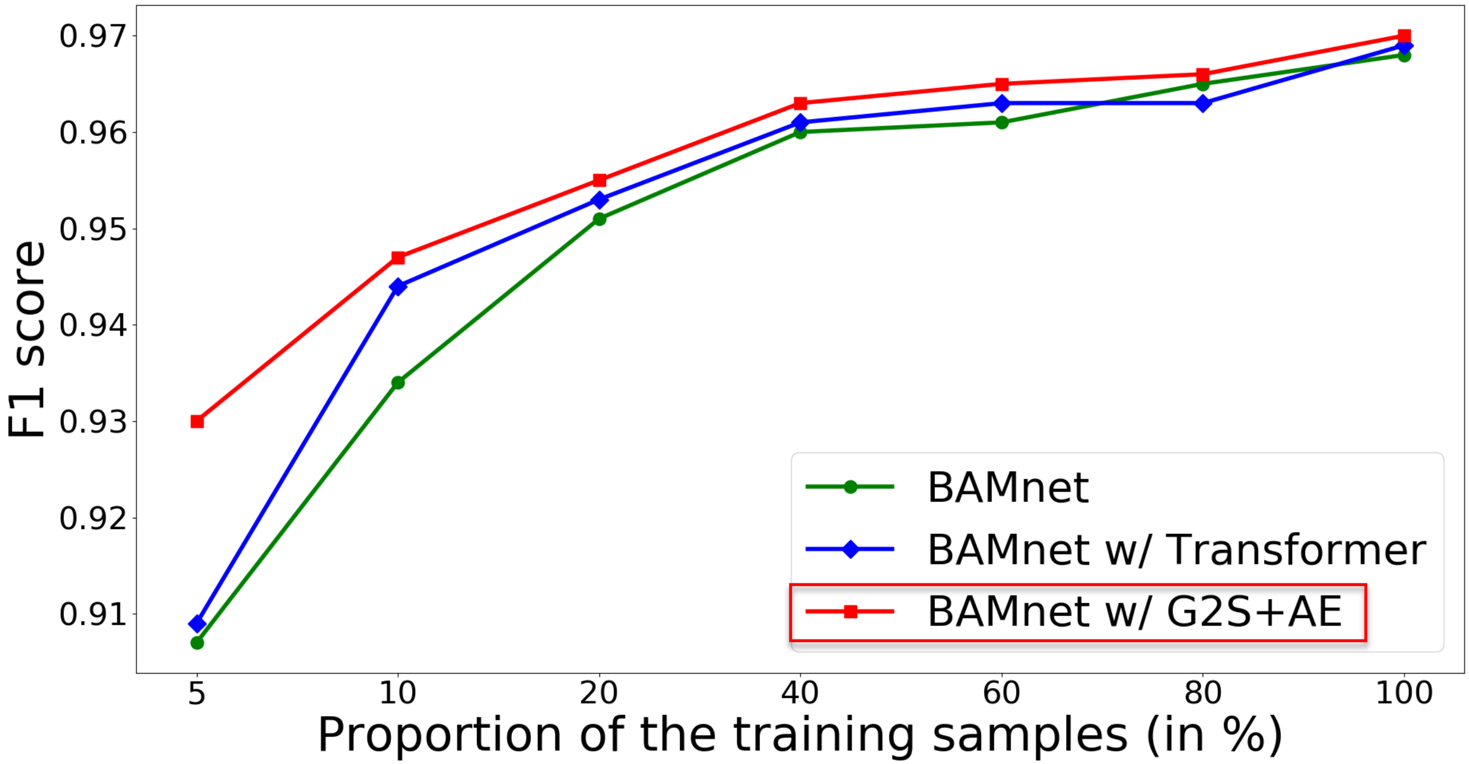


Experimental Results: Automatic Evaluation

Method	WQ			PQ		
	BLEU-4	METEOR	ROUGE-L	BLEU-4	METEOR	ROUGE-L
L2A	6.01	25.24	26.95	17.00	19.72	50.38
Transformer	8.94	13.79	32.63	56.43	43.45	73.64
MHQG+AE	11.57	29.69	35.53	25.99	33.16	58.94
G2S+AE	**29.45**	30.96	**55.45**	**61.48**	44.57	**77.72**
G2S_{edge} +AE	29.40	**31.12**	55.23	59.59	**44.70**	75.20

Table 5: Evaluation results on WQ and PQ. (higher scores indicate better results).

QG-driven Data Augmentation for QA



Outline

- Background on QA & QG
- Background on GNNs
- Dissertation Contributions
- Part I: QG from Text
- Part II: QG from KG
- Conclusion & Future Directions

Conclusion

- KBQA
  - A bidirectional attentive memory network framework for modeling the two-way flow of interactions between the questions and the KB.
- Conversational MRC
  - A Recurrent Graph Neural Network based flow mechanism for modeling the temporal dependencies in a sequence of context graphs.
- QG from KG
  - Bidirectional GNN encoder to encode the KG subgraph.
  - RNN decoder enhanced with node-level copying mechanism.
- QG from Text
  - A RL-based Graph2Seq model with a hybrid evaluator.
  - Deep Alignment Network for incorporating the answer information into the passage.

Future Directions

QA

- Complex QA
  - program induction
  - graph reasoning
- Conversational QA
- QA from multimodal data

QG

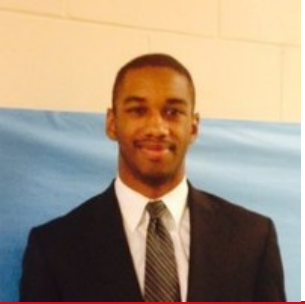
- Personalized QG
- Conversational QG
- QG from multimodal data
- Joint learning of QA & QG

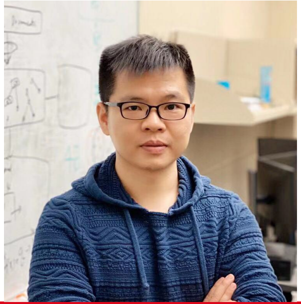









and many more!

Thank you!

Q&A

RENSSELAER POLYTECHNIC INSTITUTE
KNOWLEDGE AND THOROUGHNESS
• 1824 •

Rensselaer

why not change the world?®