YIFAN GAO

Ph.D. Candidate, Natural Language Processing, Machine Learning

+852 9794 6496 | yifangao95@gmail.com | Shatin N.T., Hong Kong SAR | Linkedin | Github | Google Scholar

EDUCATION

The Chinese University of Hong Kong — Hong Kong SAR
Ph.D. Candidate in Computer Science and Engineering — 2017 – 2021 (expected)
Research Area: Question Answering, Text Generation, Dialog Systems, Natural Language Processing
Advisors: Irwin King (IEEE Fellow, ACM Distinguished Member), Michael R. Lyu (IEEE, ACM, AAAS Fellow)
Thesis Committee: Yufei Tao (Chair), Jeffrey Xu Yu, Hsin-Hsi Chen
Ph.D. Thesis: Teaching Machine to Ask and Answer Questions: Knowledge Assessment and Information Acquisition in Reading Comprehension

Beihang University — Beijing, China
B.Eng. in Electronic and Information Engineering — 2013 – 2017
Excellent Undergraduate Award (15 out of 300), National Scholarship (6 out of 300)

Technische Universität Darmstadt — Darmstadt, Germany
Exchange Student in Computer Science — 2016 Fall
TU Darmstadt Semester Scholarship, awarded by the German Academic Exchange Service (DAAD)

EXPERIENCE

Applied Scientist Intern, Amazon Search (A9.com) — March 2021 – June 2021
Collaborators: Qingyu Yin, Zheng Li, Bing Yin — Palo Alto, CA, USA (virtually)
- Worked on multilingual keyphrase generation using Amazon dataset
- Submitted one paper to a top venue and filed one patent

Applied Scientist Intern, Amazon Web Services (AWS) AI — July 2020 – September 2020
Collaborators: Henghui Zhu, Cicero dos Santos, Zhiguo Wang, Ramesh Nallapati — New York, NY, USA (virtually)
- Investigated the query ambiguity problem in open-domain question answering
- Proposed a novel generative evidence fusion approach to find multiple interpretations for ambiguous questions
- Published one paper in ACL 2021, ranked 1st on the AmbigQA Leaderboard, and filed one patent

Research Intern, Salesforce Research — October 2019 – January 2020
Collaborators: Chien-Sheng (Jason) Wu, Shafiq Joty, Steven C.H. Hoi — Singapore
- Worked on implicative reasoning in conversational machine reading
- Proposed a new framework to explicitly track entailment states in multi-turn interactions for implicative reasoning
- Published one paper in ACL 2020, ranked 1st on the ShARC Leaderboard (Nov. 2019), and filed one patent

Research Intern, Tencent AI Lab — May 2018 – August 2019
Collaborators: Lidong Bing, Piji Li — Shenzhen, China
- Performed research on question generation in machine reading comprehension
- Published several first-author papers in conversational question generation (ACL 2019), difficulty controllable question generation (IJCAI 2019), and distractor generation (AAAI 2019)

PEER-REVIEWED PUBLICATIONS (GOOGLE SCHOLAR | DBLP | SEMANTIC SCHOLAR)

1. **Yifan Gao**, Henghui Zhu, Patrick Ng, Cicero Nogueira dos Santos, Zhiguo Wang, Feng Nan, Dejiao Zhang, Ramesh Nallapati, Andrew O. Arnold, Bing Xiang, Answering Ambiguous Questions through Generative Evidence Fusion and Round-Trip Prediction, in Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (**ACL 2021**), pages 3263–3276, August 1 - 7, 2021.

2. **Yifan Gao**, Chien-Sheng Wu, Jingjing Li, Shafiq Joty, Steven C.H. Hoi, Caiming Xiong, Irwin King, Michael R. Lyu, Discern: Discourse-Aware Entailment Reasoning Network for Conversational Machine Reading, in Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2020**), pages 2439–2449, November 16–20, 2020. [code].

3. **Yifan Gao**, Piji Li, Wei Bi, Xiaojiang Liu, Michael R. Lyu, Irwin King, Dialogue Generation on Infrequent Sentence Functions via Structured Meta-Learning, in Findings of the Association for Computational Linguistics: EMNLP 2020 (**EMNLP-Findings 2020**), pages 431–440, November 16 - 20, 2020.

4. **Yifan Gao**, Chien-Sheng Wu, Shafiq Joty, Caiming Xiong, Richard Socher, Irwin King, Michael R. Lyu, Steven C.H. Hoi, Explicit Memory Tracker with Coarse-to-Fine Reasoning for Conversational Machine Reading, in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (**ACL 2020**), pages 935–945, July 5 - 10, 2020. [code].

5. **Yifan Gao**, Piji Li, Irwin King, Michael R. Lyu, Interconnected Question Generation with Coreference Alignment and Conversation Flow Modeling, in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (**ACL 2019**), pages 4853–4862, Florence, Italy, July 28 - August 2, 2019. [code].

6. Jingjing Li*, **Yifan Gao***, Lidong Bing, Irwin King, Michael R. Lyu (* indicates equal contribution), Improving Question Generation With to the Point Context, in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (**EMNLP 2019**), pages 3216–3226, Hong Kong, China, November 3–7, 2019.

7. **Yifan Gao**, Lidong Bing, Wang Chen, Irwin King, Michael R. Lyu, Difficulty Controllable Generation of Reading Comprehension Questions, in Proceedings of the 28th International Joint Conference on Artificial Intelligence (**IJCAI 2019**). Main track. Pages 4968-4974, Macao, China, August 10-16, 2019.

8. **Yifan Gao**, Lidong Bing, Piji Li, Irwin King and Michael R. Lyu, Generating Distractors for Reading Comprehension Questions from Real Examinations, in Proceedings of the 33rd Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence (**AAAI 2019**). Main track. Pages 6423-6430, Honolulu, Hawaii, USA, January 27 - February 1, 2019. [code].

9. Wang Chen, **Yifan Gao**, Jiani Zhang, Irwin King, Michael R. Lyu, Title-Guided Encoding for Keyphrase Generation, in Proceedings of the 33rd Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence (**AAAI 2019**). Main track. Pages 6423-6430, Honolulu, Hawaii, USA, January 27 - February 1, 2019.

10. Yan Xu, Zhengyang Shen, Xin Zhang, **Yifan Gao**, Shujian Deng, Yipei Wang, Yubo Fan, Eric I-Chao Chang, Learning Multi-level Features For Sensor-based Human Action Recognition, in Pervasive and Mobile Computing 40 (2017): 324-338.

PRE-PRINTS

1. **Yifan Gao**, Jingjing Li, Michael R. Lyu, Irwin King, Open-Retrieval Conversational Machine Reading, arXiv:2102.08633, under review.

PATENTS

1. **Yifan Gao**, Chu Hong Hoi, Shafiq Rayhan Joty, and Chien-Sheng Wu, Systems and Methods for Explicit Memory Tracker with Coarse-To-Fine Reasoning in Conversational Machine Reading, US Patent, Application Number: 16/863999, Filing Date: April 30, 2020, Publication Date: June 10, 2021.

AWARDS & SCHOLARSHIPS

Global Top 100 Chinese Rising Stars in Artificial Intelligence, Baidu Scholar (2021)

Postgraduate Studentship, The Chinese University of Hong Kong (2017-2021)

National Scholarship, Ministry of Education of the P.R. China (2014)

TU Darmstadt Semester Scholarship, German Academic Exchange Service (DAAD) (2016)

PROFESSIONAL SERVICES

Journal Reviewer:
- ACM Transactions on Intelligent Systems and Technology
- Elsevier Neurocomputing
- Elsevier Neural Networks

Conference Program Committee Member:
- CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC) 2021
- Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS) 2021
- The 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2021
- International Conference on Machine Learning (ICML) 2021
- The 59th Annual Meeting of the Association for Computational Linguistics (ACL) 2021
- The 30th International Joint Conference on Artificial Intelligence (IJCAI) 2021 (Senior Program Committee)
- 2021 International Joint Conference on Neural Networks (IJCNN) 2021
- 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)
- International Conference on Learning Representations (ICLR) 2021
- The 35th AAAI Conference on Artificial Intelligence (AAAI) 2021
- The 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2020
- 2020 International Joint Conference on Neural Networks (IJCNN) 2020
- International Conference on Machine Learning (ICML) 2020
- The 58th Annual Meeting of the Association for Computational Linguistics (ACL) 2020
- The 34th AAAI Conference on Artificial Intelligence (AAAI) 2020
- The 33rd AAAI Conference on Artificial Intelligence (AAAI) 2019

MEMBERSHIP

Member, Association for Computational Linguistics (2018 – present)

Member, Association for the Advancement of Artificial Intelligence (2019)

SKILLS

Languages & Deep Learning Frameworks: Python, C/C++, Matlab, PyTorch, TensorFlow, Keras

Deep Learning Frameworks: PyTorch, TensorFlow, Keras

Developer Tools: Git, Docker, Google Cloud Platform, AWS (EC2, EFS, S3, SageMaker), PyCharm

Libraries: Huggingface Transformers, OpenNMT-py, TorchText