Tianqing Fang

Email : tfangaa@cse.ust.hk

http://www.fangtq.com

Research Interest

- Natural Language Processing
- Commonsense Knowledge Acquisition and Reasoning
- Semi/weakly-supervised Learning, Data Augmentation in NLP
- Information Extraction

Education

- **Hong Kong University of Science and Technology** — Hong Kong SAR, China
  Ph.D. of Computer Science; Supervisor: Yangqiu Song — *Aug. 2019 – Present*
- **Zhejiang University** — Hangzhou, China
  B.E. of Automation; GPA: 3.94/4.00, top 5%; Supervisor: Yang Yang — *Aug. 2015 – June. 2019*
  Minor Advanced Class of Engineering Education in Chu Ko Chen Honors College

Experience

- **University of Southern California** — Los Angeles, CA
  Visiting Research Scholar — *July 2022 - Present*
  Information extraction and commonsense reasoning
- **NVIDIA (Hong Kong)** — Hong Kong SAR, China
  Research Intern — *Feb. 2022 - June 2022*
  Semi-supervised learning on commonsense reasoning

Publications

Under Review:

[1] **Relation-aware Parameter-Efficient and Data-Efficient Commonsense Acquisition from Pre-trained Language Models**
- *Shuai Yuan, **Tianqing Fang**, Tsz Ho Chan, Yangqiu Song, Ginny Y. Wong and Simon See*
- Submitted to AAAI 2023.
  - Design relation-aware prompts and a contrastive training framework to mine *it-then* commonsense knowledge from large-scale pre-trained language models. Achieving better performance in both intrinsic and extrinsic evaluations compared with COMET.

Conference:

[2] **PseudoReasoner: Leveraging Pseudo Labels for Commonsense Knowledge Base Population**
- ***Tianqing Fang**, Quyet V. Do, Hongming Zhang, Yangqiu Song, Ginny Y. Wong and Simon See*
- Findings of EMNLP 2022.
  - Use the idea of pseudo labels to perform semi-supervised learning on CSKB Population, achieving state-of-the-art.
  - Propose a filtering strategy for pseudo labels using influence function and self distillation (the student model's own predictions).

[3] **MICO: A Multi-alternative Contrastive Learning Framework for Commonsense Knowledge Representation**
- *Ying Su, Zihao Wang, **Tianqing Fang**, Hongming Zhang, Yangqiu Song and Tong Zhang*
- Findings of EMNLP 2022.
  - A novel commonsense knowledge embedding pipeline, well used for CSKB completion and zero-shot CSQAs.

[4] **SubeventWriter: Iterative Sub-event Sequence Generation with Coherence Controller**
- *Zhaowei Wang, Hongming Zhang, **Tianqing Fang**, Yangqiu Song, Ginny Y. Wong and Simon See*
- EMNLP 2022.
  - An iterative neural text generation framework to generate multi-step instructions.

[5] **Weakly Supervised Text Classification using Supervision Signals from a Language Model**
- *Ziqian Zeng, Weimin Ni,* ***Tianqing Fang****, Xiang Li, Xinran Zhao, and Yangqiu Song.*
- Findings of Annual Conference of the North American Chapter of the Association for Computational Linguistics (Findings of NAACL). 2022.

[6] **Benchmarking Commonsense Knowledge Base Population with an Effective Evaluation Dataset**
- ***Tianqing Fang****, Weiqi Wang*, Sehyun Choi, Shibo Hao, Hongming Zhang, Yangqiu Song, Bin He*
- Conference on Empirical Methods in Natural Language Processing (**EMNLP**), 2021 (Main Conference).
  - Commonsense Knowledge Base (CSKB) Population is different from Completion as it requires reasoning over unseen assertions in external resources, while Completion only fills missing links within the CSKB.
  - Propose a dataset aligning four popular CSKBs, ConceptNet, ATOMIC, ATOMIC$^{20}_{20}$, and GLUCOSE with a large-scale eventuality graph, ASER, to populate commonsense knowledge. ˜31K triples are annotated as the evaluation set to check neural models' reasoning ability.
  - Developed KG-BertSAGE to better incorporate graph structures in the commonsense reasoning task.

[7] **DISCOS: Bridging the Gap between Discourse Knowledge and Commonsense Knowledge**
- ***Tianqing Fang****, Hongming Zhang, Weiqi Wang, Yangqiu Song, and Bin He.*
- The Web Conference (**WWW**), 2021.
  - Align the Commonsense Knowledge Base ATOMIC with a large-scale eventuality graph ASER. Use the knowledge in ATOMIC as ground-truth to train a reasoning model. Populate ATOMIC with novel edges in ASER .
  - Such commonsense knowledge acquisition method can alleviate selection bias and produce more diverse commonsense knowledge.

[8] **Do Boat and Ocean Suggest Beach? Dialogue Summarization with External Knowledge**
- ***Tianqing Fang****, Haojie Pan, Hongming Zhang, Yangqiu Song, Kun Xu, Dong Yu.*
- Conference on Automated Knowledge Base Construction (**AKBC**). 2021.
  - Address the situation where summarization may include something out of the dialogue context but can be implicitly inferred. Develop a knowledge-attention network to tackle this problem and achieves promising results.

[9] **Probing Toxic Content in Large Pre-Trained Language Models**
Nedjma Ousidhoum, Xinran Zhao, ***Tianqing Fang****, Yangqiu Song, and Dit-Yan Yeung*
Annual Meeting of the Association for Computational Linguistics (**ACL**). 2021.

Journal:

[10] **ASER: Towards Large-scale Commonsense Knowledge Acquisition via Higher-order Selectional Preference over Eventualities**
Hongming Zhang, Xin Liu*, Haojie Pan*, Haowen Ke, Jiefu Ou,* ***Tianqing Fang****, and Yangqiu Song.*
Artificial Intelligence. 2022

Preprint:

[11] **Acquiring and Modelling Abstract Commonsense Knowledge via Conceptualization**
Mutian He, ***Tianqing Fang****, Weiqi Wang, and Yangqiu Song.*
Submitted to Journal of Artificial Intelligence. arxiv.2206.01532, 2022

Academic Achievements

- HKUST RedBird Academic Excellence Award for Continuing PhD Students in 2021/22 (2022)
- Hong Kong Ph.D. Fellowship (2019-2023)
- Special Scholarship for Undergraduate Students in Zhejiang University (One of the highest awards for undergraduates) (2018)
- 1st Place and MATLAB Innovation Award (1st/36k+) in Contemporary Undergraduate Mathematical Contest in Modeling (The most authoritative mathematical modeling competition in China) (2017)
- National Scholarship (top 3%, ZJU, 2016)

Skills

- **Programming skills**: C++, Python
- **Languages**: English (TOEFL 110, 26 in speaking), Mandarin Chinese (Native).
- **Miscs**: I enjoy taking pictures. Street scenery is my favorite.