

# Siru Ouyang

3-220 SEIEE Building – Shanghai Jiao Tong University – Shanghai, China

☎ +86 15869677967 • ✉ oysr0926@gmail.com • 🌐 ozyyshr.github.io

## Education

---

### Shanghai Jiao Tong University

Shanghai, China

Sep. 2018 - Jun. 2022

- Senior undergraduate, Department of Computer Science (IEEE honor class)
- GPA: Overall: 89.2/100 | Major: 90.7/100 | Final Year: 90.5/100
- Standard Test: TOEFL: 107 (R29, L28, S25, W25), GRE: V153, Q170, W4.0
- Advisor: Professor [Hai Zhao](#), Professor [Xiaofeng Gao](#)

### Georgia Institute of Technology

Remote

June 2021 - present

- Research Intern, School of Interactive Computing
- Advisor: Professor [Diyi Yang](#)

## Papers

---

### Dialogue Graph Modeling for Conversational Machine Reading

- [Siru Ouyang\\*](#), [Zhuosheng Zhang\\*](#), [Hai Zhao](#).
- In *Findings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*. **ACL-IJCNLP 2021**.

### Smoothing Dialogue States for Open Conversational Machine Reading

- [Zhuosheng Zhang\\*](#), [Siru Ouyang](#), [Hai Zhao](#), [Masao Utiyama](#), [Eiichiro Sumita](#).
- In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. **EMNLP 2021**.

### Two-Hop Relay Deployment Based on User Trajectory in Wireless Networks

- [Zhiyao Li](#), [Siru Ouyang](#), [Xiaofeng Gao](#) and [Guihai Chen](#).
- Accepted by *the Computer Journal* 2021.

### Fact-driven Logical Reasoning

- [Siru Ouyang\\*](#), [Zhuosheng Zhang\\*](#), [Hai Zhao](#).
- In submission for *the International Conference on Learning Representations*. **ICLR 2022**.

### Logic-aware Pre-Training for Language Models

- [Siru Ouyang\\*](#), [Zhuosheng Zhang\\*](#), [Hai Zhao](#).
- In submission for *the International Conference on Learning Representations*. **ICLR 2022**.

### Compositional Data Augmentation for Abstractive Conversation Summarization

- [Siru Ouyang](#), [Jiaao Chen](#), [Diyi Yang](#).
- In submission for *the Annual Meeting of the Association for Computational Linguistics* **ACL 2022**.

## Honors and Awards

---

- **Rongchang Scientific and Technological Innovation Scholarship** (RMB ¥30,000) (Rank 1<sup>st</sup> in the School of Electronic Information and Electrical Engineering), Shanghai Jiao Tong University. 2021
- **SenseTime Scholarship** (RMB ¥20,000) (31 out of all students for artificial intelligence in China) SenseTime Incorporation. 2021
- **Class-B Scholarship in Campus** (5% of all the students in the School of Electronic Information and Electrical Engineering), Shanghai Jiao Tong University. 2019 and 2020
- **Google Women Tech Makers Scholarship** (110 out of 2800 students around the whole Pacific Rim), Google. 2020

## Research Experiences

---

### Compositional Data Augmentation for Dialogue Summarization

Advisor: Professor Diyi Yang, SIC, GIT

Jul. 2021 - Nov. 2021

- Proposed a simple yet effective compositional data augmentation method, COMPO, for generating diverse and high-quality pairs of conversations and summaries based on conversation structures.
- Introduced model-level distillation algorithms to learn concise representations from a teacher model to deal with noise.
- Outperformed prior SOTA baselines by a large margin in terms of qualitative and quantitative evaluation.
- In preparation as the first-author for [ACL 2022 \[pdf\]](#).

### Logic Pre-Training of Language Models

Advisor: Professor Hai Zhao, CSD, SJTU

Jun. 2021 - Sep. 2021

- Proposed a novel pre-training framework PROPHET including a newly introduced knowledge basis *fact* and pre-training objectives for capturing logic relations in challenging NLU tasks.
- Outperformed strong baselines by a large margin on GLUE benchmark (NLU), ReClor (Logical Reasoning) and DocRE (Relation Extraction).
- In submission as a full paper at [ICLR 2022](#). [[pdf](#)]

### Logical Reasoning in Natural Language Understanding

Advisor: Professor Hai Zhao, CSD, SJTU

Mar. 2021 - May. 2021

- Proposed to extract a broad *Fact Unit* to effectively cover indispensable logic reasoning basis representing the relations of "who-did-what-to-whom" or "who-is-what".
- Designed FOCAL REASONER which builds super-graphs on top of fact units to capture both global connections and local concepts inside the fact.
- Achieved the new state-of-the-art results on ReClor and LogiQA.
- In submission as a full paper at [ICLR 2022](#). [[pdf](#)]

### End-to-End Open-Domain Conversational Comprehension

Advisor: Professor Hai Zhao, CSD, SJTU

Feb. 2021 - Apr. 2021

- Investigated the open-retrieval setting for conversational machine reading.
- Designed an end-to-end framework where the dialogue states for decision making are employed for question generation, in contrast to the independent models or pipeline systems in previous studies for CMR task.
- Outperformed strong baselines in both the ShARC and OrShARC datasets.
- Accepted as the co-first author by [EMNLP 2021](#).

### Dialogue Graph Modeling for Reading Comprehension

Advisor: Professor Hai Zhao, CSD, SJTU

Nov. 2020 - Jan. 2021

- Proposed a framework DGM for conversational machine reading to model discourse structure and relations which has not been considered by previous approaches.
- Designed an implicit and explicit discourse graph to analyze the complex rule document and their interactions with user scenario.
- Achieved the new state-of-the-art results on ShARC benchmark.
- Accepted as the first author at [ACL Findings 2021](#).

### Relay Deployment in Wireless Networks

Advisor: Professor Xiaofeng Gao, CSD, SJTU

May. 2020 - Jul. 2020

- Solve the traditional relay deployment problem for unstationary user movements.
- Proposed the concept "Demand Nodes" representing the locations where users frequently pass or stay, and convert the problem into a Demand Node Coverage (DNC) problem, which proved to be NP-complete.
- Designed an approximation algorithm to solve DNC problem, and simulated the method on five real-world trajectory datasets in CRAWDAD, and results demonstrate that the proposed algorithm can obtain high coverage for users in motion, which leads to better user experience.
- Accepted as a full paper in [The Computer Journal](#).

### Selected Academic Presentations and Open Source Projects

- [Dialogue Graph Modeling for Conversational Machine Reading](#). ACL 2021 Conference, Online.
- [Implementation of DGM](#). Discourse-based document comprehension.