

## EDUCATION

### University of British Columbia

*PhD in Computer Science • Department of Computer Science*

Vancouver, BC, Canada  
(Expected) Sep 2025 – Aug 2029

- **Advisors:** Profs. [Hila Gonen](#), [Jian Zhu](#) and [Isabel Papadimitriou](#)
- **Research Foci:** Multilingual & Multicultural LLMs/Speech Models/VLMs; Interpretability

### ETH Zurich (Swiss Federal Institute of Technology Zurich)

*MSc in Data Science • Department of Computer Science (D-InfK)*

Zurich, ZH, Switzerland  
Sep 2021 – Dec 2024

- **Advisor:** Profs. [Ryan Cotterell](#) and [Mrinmaya Sachan](#) **Research Foci:** NLP, Reasoning Ability of LLMs
- **GPA:** 5.4/6.0 (Top 30% of Cohort) **Award:** ETH Scholarship

### The Chinese University of Hong Kong, Shenzhen

*BSc in Statistics • Stream in Data Science • School of Science and Engineering*

Shenzhen, GD, China & Hong Kong  
Sep 2017 – Nov 2020

- First Class Honours. Completed the 4-year program one year in advance.
- **Advisor:** Prof. [Jianfeng Mao](#) **Research Foci:** AI/ML Algorithm Applications
- **Cumulative GPA:** 3.764/4.0 (Rank 11/313 of School) **Major GPA:** 3.847/4.0 (Rank 5/150 of Statistics Cohort)
- **Awards:** Academic Performance Scholarship; Dean's List; Master's List (All 3 consecutive years)

### University of California, Berkeley

*Summer Session • Non-degree • GPA: 4.0/4.0*

Berkeley, CA, United States  
Jun 2019 – Aug 2019

## RESEARCH EXPERIENCE

### CompLING Lab, University of Waterloo & Vector Institute

*Visiting Scholar • Supervisor: Prof. [Freda Shi](#)*

Waterloo, ON, Canada  
Sep 2024 – Feb 2025

- Analyzed systematically multilingual in-context learning; Showed that demonstrations in mixed high-resource languages consistently outperform English-only ones across the board, particularly for tasks written in low-resource languages [1].
- Studied that the presence of irrelevant non-English sentences in the prompt yields measurable gains, suggesting the effectiveness of multilingual exposure itself.
- Highlighted the potential of leveraging multilingual resources to bridge the performance gap for underrepresented languages.

### LRE Lab, ETH Zurich

*Research Assistant • Supervisor: Prof. [Mrinmaya Sachan](#)*

Zurich, ZH, Switzerland  
May 2024 – Aug 2024

- Found that VLMs can nearly perfectly identify entities in diagrams and demonstrate strong reasoning abilities on entities, while barely identify relations, and they are unable to reason about them [2].
- Validated that VLMs rely on knowledge shortcuts answering complex relation reasoning questions, indicating that the successes of VLMs on diagram reasoning tasks give a false illusion of their diagram understanding capabilities.
- Curated a test suite of 1k real diagrams and 500 synthetic diagrams with 8 annotated questions each to benchmark VLMs' ability to recognize and reason about entities and relations with/without background knowledge.

### Rycolab, ETH Zurich

*Research Assistant • Supervisor: Prof. [Ryan Cotterell](#)*

Zurich, ZH, Switzerland  
Jun 2023 – Dec 2023

- Investigated the probing accuracy of probes across varying complexities on pretrained and randomized representations (rep), found that more complex probes perform worse, especially when linguistic information is not present in the representations [4].
- Revealed a counter-intuitive trend that increased complexity leads to lower accuracy and that the noise in randomized rep hampers learning, while for pretrained rep, high and stable probing accuracy is observed, insensitive to hyperparam variations.

### Chair of Applied Economics, ETH Zurich

*Research Assistant • Part-time • Supervisor: Prof. [Peter Egger](#), Dr. [Susie Rao](#)*

Zurich, ZH, Switzerland  
Jan 2023 – Aug 2024

- Built a chatbot combining Llama3 and SemOpenAlex knowledge graph to support meta-science research, recommending similar papers and authors and answering users' questions about scientific literatures [3].
- Applied LLMs for classification and keyword extraction of literatures; Developed an open-source and cross-discipline scientific annotation and inference engine demo that allows user to annotate features easily and to fine-tune downstream tasks in meta-science studies [5].

### Shenzhen Research Institute of Big Data & CUHK-Shenzhen

*Research Assistant • Supervisor: Prof. [Jianfeng Mao](#), Dr. [Yuan Wang](#)*

Shenzhen, GD, China  
Dec 2020 – May 2021

- Developed a three-staged online decision-making framework with colleagues to optimize sector-wise air traffic flow with multi-agent reinforcement learning and combinatorial programming [6].
- Developed of a new two-stage medium-term data-driven model that combined various algorithms such as ARIMA, DBSCAN, LSTM and random forest, to predict the estimated time of arrival in a multi-airport system [7].

## INTERNSHIPS

---

### Tencent

Algorithm Engineer • Financial Technology (FiT) • Supervisor: Liang Chen

Shenzhen, GD, China

Jul 2021 – Sep 2021

- Collaborated with colleagues to propose a new algo-trading based on neural models (XGBoost and LSTM) and modern portfolio theories (CAPM, Black-Litterman). Backtested the strategy and achieved a 12% outperformance over S&P 500.
- Collected and analyzed US stock indicators and designed their calculation methods or alternatives in China's stock market.

### Huawei

Software Development Engineer • Intelligent Automotive Solution (IAS)

Shenzhen, GD, China

Jun 2020 – Aug 2020

- Developed audit-log module by Golang to concurrently monitor and report crucial behaviors of the administrators and users.
- Completed a set of APIs' design, development and deployment of the IAS cloud identity management system.
- Wrote white- and black-box test cases, and undertook daily work of deploying and optimizing automated test cases.

## PUBLICATIONS & PREPRINTS

---

- [1] [Yilei Tu](#), Andrew Xue, and Freda Shi. [Blessing of Multilinguality: A Systematic Analysis of Multilingual In-Context Learning](#). arXiv preprint: arXiv:2502.11364.
- [2] Yifan Hou\*, Buse Giledereli\*, [Yilei Tu\\*](#), and Mrinmaya Sachan. [Do Vision-Language Models Really Understand Visual Language?](#) arXiv preprint: arXiv:2410.00193.
- [3] [Yilei Tu\\*](#), Noah Mamié\*, Xi Susie Rao, and Peter Egger. [Knowledge-Enhanced Academic Chatbot: Harnessing Large Language Models and Knowledge Graphs](#). Under revision.
- [4] [Yilei Tu](#), Jiaoda Li, and Ryan Cotterell. [Complex Probes are Favored: A Revisit of Probe Complexity](#). Under revision.
- [5] Susie Xi Rao, [Yilei Tu](#), and Peter H. Egger. [SAINE: Scientific Annotation and Inference Engine of Scientific Research](#). In *IJCNLP-AACL 2023: System Demonstrations*, pages 41–58, Bali, Indonesia, November 2023. Association for Computational Linguistics.
- [6] Yuan Wang, Weilin Cai, [Yilei Tu](#), and Jianfeng Mao. [Reinforcement-Learning-Informed Prescriptive Analytics for Air Traffic Flow Management](#). *IEEE Transactions on Automation Science and Engineering*, pages 1–15, 2023.
- [7] Lechen Wang, Jianfeng Mao, Lishuai Li, Xuechun Li, and [Yilei Tu](#). [Prediction of estimated time of arrival for multi-airport systems via “Bubble” mechanism](#). *Transportation Research Part C: Emerging Technologies*, 149:104065, 2023.

## TEACHING

---

### Department of Mathematics, ETH Zurich

Teaching Assistant • Part-time • Institute for Operations Research

Zurich, ZH, Switzerland

Sep 2023 – Jan 2024

- 401-0647-00L Introduction to Mathematical Optimization; Lecturer: Dr. David Adjashvili

## EXTRACURRICULAR

---

### Board Game Club (BGC), CUHK-Shenzhen

Co-founder • Minister of Publicity • Excellent Club Organization Prize

Shenzhen, GD, China

Oct 2017 – Sep 2019

- Coordinated with other departments to organize monthly board game events of 200+ players.
- Managed social media accounts; Promote BGC activities and cooperate with other board game organizations.

## SKILLS

---

**Programming Languages & Tools:** Python (PyTorch, HuggingFace, vLLM), R, Golang, SQL, Git, Unix, JetBrains, Dokcer, Tex  
**Natural Languages:** English (Advanced/C1, TOEFL 109); Mandarin Chinese & Sichuanese Dialects (Native); Cantonese (Fluent); Standard High German (A2); Japanese (N4)

## 外部链接 EXTERNAL LINKS

---

所有文本和图标若颜色为军校蓝则包含超链接。若 PDF 阅读器不支持点击超链接，下面提供了具体的 URL。  
In case the PDF reader does not support the clicking of links, some URLs are provided explicitly below.

- 个人网站 Personal Website: <https://yileitu.github.io/>
- 谷歌学术 Google Scholar: <https://scholar.google.com/citations?user=jPeyGn4AAAAJ&hl=en>
- GitHub: <https://github.com/yileitu>
- 领英 LinkedIn: <https://www.linkedin.com/in/yileitu/>
- 论文 Paper [1]: <https://arxiv.org/abs/2502.11364>
- 论文 Paper [2]: <https://www.arxiv.org/abs/2410.00193>
- 论文 Paper [3]: <https://yileitu.github.io/assets/pdf/Chatbot.pdf>
- 论文 Paper [4]: [https://yileitu.github.io/assets/pdf/EACL\\_submission.pdf](https://yileitu.github.io/assets/pdf/EACL_submission.pdf)
- 论文 Paper [5]: <https://www.arxiv.org/abs/2302.14468>
- 论文 Paper [6]: <https://ieeexplore.ieee.org/abstract/document/10258410>
- 论文 Paper [7]: <https://www.sciencedirect.com/science/article/pii/S0968090X23000542>