Xinyi Wang

xinyi_wang@ucsb.edu
Homepage: https://wangxinyilinda.github.io/
Tel.: (226)387-9013

About Me

I am a forth-year Ph.D. candidate in the computer science department at the University of California, Santa Barbara. I am interested in understanding deep learning models, especially pre-trained large language models, using principled causality-based/probabilistic approaches.

Education

- **University of California, Santa Barbara (UCSB)** — Santa Barbara, CA, US
  Ph.D. in Computer Science (expected) — *09.2020 - present*
  - GPA: 4.0/4.0
  - Advisor: William Yang Wang
  - Honors and Awards: Academic Excellence Fellowship (2020)

- **The Hong Kong University of Science and Technology (HKUST)** — Hong Kong, China
  B.Sc in Applied Mathematics and Computer Science — *09.2016 - 07.2020*
  - GPA: 3.7/4.3
  - Honors and Awards: Chern Class Talent Scholarship (2017 - 2020), University's Scholarship Scheme for Continuing Undergraduate Students (2017 - 2020), HKSAR Government Scholarship Fund - Reaching Out Award (2019 - 2020), Chern Class Achievement Scholarship (2020), The 15th Epsilon Fund Award (2020), Joseph Needham Merit Scholarship (2020)

- **University of California, Los Angeles (UCLA)** — Los Angeles, CA, US
  Term exchange in Mathematics (Non-degree) — *09.2019 - 12.2019*
  - GPA: 3.9/4.0 (Dean's Honors List)

Publications

- **Xinyi Wang**, Alfonso Amayuelas, Kexun Zhang, Liangming Pan, Wenhu Chen, William Yang Wang. *Understanding the Reasoning Ability of Language Models From the Perspective of Reasoning Paths Aggregation.* Arxiv preprint 2024. [paper]
- **Xinyi Wang**, Lucas Caccia, Oleksiy Ostapenko, Xingdi Yuan, William Yang Wang, Alessandro Sordoni. *Guiding Language Model Math Reasoning with Planning Tokens.* Arxiv preprint 2023. [paper]
- Liangming Pan, Michael Saxon, Wenda Xu, Deepak Nathani, **Xinyi Wang**, William Yang Wang. *Automatically Correcting Large Language Models: Surveying the landscape of diverse self-correction strategies.* TACL 2023. [paper]
- Liangming Pan, Alon Albalak, **Xinyi Wang**, William Yang Wang. *Logic-LM: Empowering Large Language Models with Symbolic Solvers for Faithful Logical Reasoning.* Findings of EMNLP 2023. [paper]
- Wenhu Chen, Ming Yin, Max Ku, Pan Lu, Yixin Wan, Xueguang Ma, Jianyu Xu, **Xinyi Wang**, Tony Xia. *TheoremQA: A Theorem-driven Question Answering dataset.* EMNLP 2023. [paper]
- Wanrong Zhu, **Xinyi Wang**, Yujie Lu, Tsu-Jui Fu, Xin Eric Wang, Miguel Eckstein, William Yang Wang. *Collaborative Generative AI: Integrating GPT-k for Efficient Editing in Text-to-Image Generation.* EMNLP 2023. [paper]
- **Xinyi Wang**, Wanrong Zhu, William Wang. *Large Language Models Are Latent Variable Models: Explaining and Finding Good Demonstrations for In-Context Learning.* NeurIPS 2023, poster. [paper]
- Wenhu Chen, Xueguang Ma, **Xinyi Wang**, William W. Cohen. *Program of Thoughts Prompting: Disentangling Computation from Reasoning for Numerical Reasoning Tasks.* TMLR 2023. [paper]

- **Xinyi Wang**, Michael Saxon, Jiachen Li, Hongyang Zhang, Kun Zhang, William Yang Wang. *Causal Balancing for Domain Generalization*. ICLR 2023, poster. [paper]
- Michael Saxon, **Xinyi Wang**, Wenda Xu, William Yang Wang. *Relation Leakage in Elicited Natural Language Inference Datasets*. EACL 2023. [paper]
- Wenhu Chen, **Xinyi Wang**, William Yang Wang. *A Dataset for Answering Time-Sensitive Questions*. NeurIPS 2021 Datasets and Benchmarks Track, poster. [paper]
- **Xinyi Wang**, Wenhu Chen, Michael Saxon, William Yang Wang. *Counterfactual Maximum Likelihood Estimation for Training Deep Networks*. NeurIPS 2021, poster. [paper]
- Michael Saxon, Sharon Levy, **Xinyi Wang**, Alon Albalak, William Yang Wang. *Modeling Disclosive Transparency in NLP Application Descriptions*. EMNLP 2021, oral. [paper]
- **Xinyi Wang***, Haiqin Yang*, Liang Zhao, Yang Mo and Jianping Shen. *RefBERT: Compressing BERT by Referencing to Pre-computed Representations*. IJCNN 2021, oral. [1] [paper]
- **Xinyi Wang**, Yi Yang. *Neural Topic Model with Attention for Supervised Learning*. AISTATS 2020, poster. [paper]

Research Experience

- **Research Intern at Microsoft Research** — Montreal, QC, Canada
  Mentor: Alessandro Sordoni — *06.2023 - 10.2023*
  - Topic: parameter-efficient fine-tuning to improve math reasoning ability of large language models.
- **Graduate Student Researcher at UCSB** — Santa Barbara, CA, US
  Mentor: William Yang Wang — *09.2020 - present*
  - Topic: make better use of large language models by developing a principled interpretation of them.
- **Assistant Algorithm Engineer at PingAn AI** — ShenZhen, China
  Mentor: Haiqin Yang — *06.2020 - 09.2020*
  - Topic: retrieval augmented BERT distillation.
- **Research Assistant at HKUST** — Hong Kong, China
  Mentor: Yi Yang — *09.2018 - 09.2019*
  - Topic: word embedding with financial knowledge, supervised neural topic model with attention.

Services

- 2021 Program Committee: NeurIPS Datasets and Benchmarks Track
- 2022 Program Committee: AAAI
- 2023 Program Committee: NeurIPS, AAAI
- 2024 Program Committee: ICLR, ICML

[1] * denotes equal contribution.