Yucen Li (Lily)

yucenli@gmail.com | yucenli.com

Education

Carnegie Mellon University Aug 2015 – Dec 2018

- B.S. in Computer Science, 3.83/4.0 GPA
- Minor in Statistics, Minor in Language Technologies

Research Experience

Newtonian Monte Carlo Aug 2019 – Present

Facebook Probabilistic Programming Languages

- Developed second order gradient-based Markov Chain Monte Carlo algorithm
- Explored behavior on different models and curvatures
- Improved performance through step size learning rate and covariance scaling

Bean Machine Aug 2019 – Present

Facebook Probabilistic Programming Languages

- Developed a declarative probabilistic programming language with explicit dependencies
- Enabled users to program inference techniques using custom proposers
- Implemented proposal algorithms such as HMC and NUTS for single-site inference

Signals to Contributers in Open-Source Projects Aug 2018 – May 2019

Socio-Technical Research Using Data Excavation Lab

- Conducted analysis on GitHub READMEs to identify similarities between projects
- Mined GitHub repository data to quantify signals for new contributors
- Modeled the number of newcomers as a function of the signals to determine significance

Multi-Word Expressions in Word Embeddings Aug 2018 – Dec 2018

Linguistics Lab

- Implemented techniques for automatically identifying multi-word expressions
- Analyzed granularity of words in a variety of languages
- Evaluated cross-lingual embeddings with multi-word expressions

Cross-lingual Dependency Parsing Nov 2017 – May 2018

Linguistics Lab

- Analyzed linguistic typology of languages such as subject word order
- Evaluated which language features are most relevant for cross-lingual dependency parsing
- Configured multilingual model for experiments with different combinations of languages

Teaching

Principles of Software Engineering Spring 2019

Teaching Assistant

- Led weekly recitations, held office hours, and graded homeworks

Concepts in Mathematics Spring 2017 – Fall 2018

Academic Development Leader

- Worked with professors to design curriculum and problem sets
- Led weekly review groups focused on collaborative learning
- Facilitated other leaders in leading successful sessions for students

Matrices and Linear Transformations Fall 2016

Academic Development Leader

- Worked with professors to lead collaborative learning groups

Work Experience

Facebook Aug 2019 – Present

Software Engineer

- Developed Bean Machine and MCMC inference algorithms
- Collaborated with Marketing Science team to use Bean Machine to predict brand lift

Facebook May 2018 – Aug 2018

Software Engineering Intern

- Worked on Instagram Explore product team to add functionality to tag users in videos
- Implemented reusable component for profile which is now standard across app

Facebook May 2017 – Aug 2017

Software Engineering Intern

- Worked on the Mobile Interface Health team to classify HTTP requests on the Android app
- Extended parsing in Duckling, an open-source Haskell text library, to include weights

Hyland May 2016 – Aug 2016

Software Engineering Intern

- Optimized Microsoft Word placeholders through Microsoft Word plugin using C# and WPF
- Used .NET framework to design controls for bulk creation of placeholders

Publications

1. Nazanin Tehrani, Nimar S. Arora, **Yucen Lily Li**, Kinjal Divesh Shah, David Noursi, Michael Tingley, Narjes Torabi, Sepehr Masouleh, Eric Lippert, and Erik Meijer. Bean machine: A declarative probabilistic programming language for efficient programmable inference. In *International Conference on Probabilistic Graphical Models (PGM)*, 2020

2. Huilian Sophie Qiu, **Yucen Lily Li**, Susmita Padala, Anita Sarma, and Bogdan Vasilescu. The signals that potential contributors look for when choosing open-source projects. In *Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*, 2019

3. Naoki Otani, Satoru Ozaki, Xingyuan Zhao, **Yucen Lily Li**, Micaelah St. Johns, and Lori Levin. Pre-tokenization of multi-word expressions in cross-lingual word embeddings. *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020

4. Feynman Liang, Nimar Arora, Nazanin Tehrani, **Yucen Lily Li**, Michael Tingley, and Erik Meijer. Accelerating metropolis-hastings with lightweight inference compilation. *arXiv preprint arXiv:2010.12128*, 2020

5. Nimar S. Arora, Nazanin Khosravani Tehrani, Kinjal Divesh Shah, Michael Tingley, **Yucen Lily Li**, Narjes Torabi, David Noursi, Sepehr Akhavan Masouleh, Eric Lippert, and Erik Meijer. Newtonian monte carlo: single-site mcmc meets second-order gradient methods. *arXiv preprint arXiv:2001.05567*, 2020

6. Sourabh Kulkarni, Kinjal Divesh Shah, Nimar Arora, Xiaoyan Wang, **Yucen Lily Li**, Nazanin Khosravani Tehrani, Michael Tingley, David Noursi, Narjes Torabi, Sepehr Akhavan Masouleh, Eric Lippert, and Erik Meijer. Ppl bench: Evaluation framework for probabilistic programming languages. *arXiv preprint arXiv:2010.08886*, 2020