Hebi Li

lihebi.com @lihebi hebi-li-395989a2 lihebi.com@gmail.com +1(515)708-8131

Education

- **Iowa State University** *Aug. 2014 - Aug. 2022 (expected)*
  PhD Candidate in Dept. of Computer Science. *Advisor: Prof. Jin Tian*
  Research Topics: Artificial Intelligence & Machine Learning; Programming Language Design
- **University of Science and Technology of China** *Sep. 2010 - Jun. 2014*
  B.E. in Electrical Engineering

Professional Skills

- **Programming**: Python, Julia, Javascript, Racket/Scheme/Lisp, Meta Prog, DSL, C/C++, Clang/LLVM
- **System & Web**: Linux, Container, Docker, Tmux, Jupyter, ReactJS, ExpressJS, GraphQL, SQL, NoSQL
- **Machine Learning**: Tensorflow, Pytorch, Flux, MLP, CNN, LSTM, Auto-encoders, Scikit-learn, Numpy, Pandas

Work Experience

- **ByteDance Inc.**
  PhD Research Intern. Working on Graph Neural Networks. In progress. *Jan. 2022 - May. 2022*
- **Iowa State University**
  Research Assistant. Working on Artificial Intelligence. *May. 2015 - Dec. 2021*

Research Projects

- **Self-supervised Causal DAG Structure Learning [1]** *Nov. 2019 - Aug. 2021*
  - Proposed the first self-supervised learning formulation of the causal graph DAG structure learning problem.
  - Achieved state-of-the-art accuracy, and order-of-magnitude faster inference comparing with the literature.
  - Tech stack: Julia, Python & Tensorflow, Neural Networks
  - Links: `https://github.com/lihebi/DAG-EQ`, `https://arxiv.org/abs/2006.04697`

- **Embedded Programming Language and System for PCB Designs [2]** *Feb. 2020 - Oct. 2021*
  - Designed the first open-source embedded domain-specific language (eDSL) for creating PCBs
  - Implemented an end-to-end workflow from computer program to ready-to-manufacture placed & routed PCB
  - Published as conference paper at IEEE/ACM **DAC-2021** .
  - Links: `https://bhdl.org`, `https://github.com/bhdl/bhdl`, `http://lihebi.com/assets/BHDL.pdf`

- **AI-driven automatic PCB routing algorithm [3]** *Feb. 2020 - Oct. 2021*
  - Designed DRL-MCTS, an AI-driven automatic routing algorithm for PCBs using deep reinforcement learning (RL) and Monte Carlo Tree Search (MCTS). It yields 33.3% higher success rate than A* and PPO.
  - Accepted as a conference paper at **VLSI-DAT 2022** .

- **Group-wise Feature Selection for Supervised Learning [4]** *May. 2020 - Jan. 2022*
  - Proposed a novel kind of feature selection problem with a focus on groupness.
  - Utilize the Mixture of Experts (MoE) model with Gumbel-Softmax to learn the group membership and feature selector simultaneously; achieve much better performance than related baselines.
  - Accepted as a conference paper at **ICASSP 2022**

- **End-to-end Semantics-based Summary Assessment [5]** *Feb. 2019 - Feb. 2020*
  - Designed a reference-free supervised approach BetterRouge for Summary Quality Assessment.
  - Proposed two negative sampling methods for fully automatic training data augmentation.
  - Utilized MLP, CNN and LSTM, and applied the model to both word embedding and sentence embedding.
  - Our approach significantly outperforms reference-free baselines for extractive summarizers.
  - Link: `https://arxiv.org/abs/2005.06377`

- **Hierarchical IDE for Interactive Development at Scale (WIP) [6]** *Aug. 2020 - Present*
  - Designed and developed **CodePod**, the first scalable Jupyter-like Integrated Dev Env (IDE) for interactive development. Thanks to its novel namespace rules, CodePod permits developing production-scale codebase.
  - CodePod is implemented as a fullstack web app using ReactJS, ExpressJS, GraphQL, PostgreSQL, Prisma ORM.

- **Purifying Adversarial Defense with Adversarial Auto-encoders [7]** *Aug. 2018 - Sep. 2019*
  - Proposed a novel adversarial auto-encoder AdvAE that protects Neural Networks from adversarial attacks.
  - Showed that AdvAE significantly outperforms other purifying-based adversarial defense techniques.
  - Links: `https://github.com/lihebi/AdvAE`, `https://arxiv.org/abs/1905.10729`

- **Demand-driven Dynamic Program Analysis via Syntactic Patching**   *May. 2015 - Mar. 2018*
    - Designed the first algorithm that constructs minimal executable C++ sub-programs
    - Implement with Clang/LLVM compiler framework, open source at `https://github.com/lihebi/Helium`
- **USTC RoboGame-2012**   *May. 2012 - Aug. 2012*
    - Designed a robot car that fetches flowers from point A to point B. Notable features: identifying and picking up the flowers; auto-routing with computer vision; climbing up/down the stairs
    - Led a team of 4 and ranked 3rd among 26 teams. `https://youtu.be/NOEbvINeiy4`

Publications

[1] **Hebi Li**, Qi Xiao, and Jin Tian. *"Supervised Whole DAG Causal Discovery."*, arXiv:2006.04697 (2020). full paper in submission Source Code: `https://github.com/lihebi/DAG-EQ`

[2] **Hebi Li**, Y. He, Qi Xiao, Jin Tian, F. Bao. *"BHDL: A Lucid, Expressive, and Embedded Programming Language and System for PCB Designs"*, published as a conference paper at IEEE/ACM **DAC 2021**. web: `https://bhdl.org`, source: `github.com/bhdl/bhdl`

[3] Youbiao He, **Hebi Li**, Forrest Bao, Jin Tian, *"Circuit Routing Using Monte Carlo Tree Search and Deep Reinforcement Learning"*, arXiv preprint arXiv:2006.13607 (2020). published as a conference paper at **VLSI-DAT 2022**.

[4] Qi Xiao, **Hebi Li**, J. Tian and Z. Wang. *"Groupwise Feature Selection for Supervised Learning"*, published as a conference paper at **ICASSP 2022**

[5] F. Bao, **Hebi Li**, Ge Luo, Cen Chen, Y. Yang, and M. Qiu. *"End-to-end Semantics-based Summary Quality Assessment for Single-document Summarization."* preprint arXiv:2005.06377 (2020), full paper in submission.

[6] **Hebi Li**, F. Bao, Qi Xiao and Jin Tian, *"CodePod: A Hierarchical Jupyter for Interactive Development at Scale"*, in submission

[7] **Hebi Li**, Qi Xiao, Shixin Tian, and Jin Tian. *"Purifying Adversarial Perturbation with Adversarially Trained Auto-encoders."*, preprint arXiv:1905.10729 (2019). Source Code: `https://github.com/lihebi/AdvAE`