Hebi Li

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

Education

- **Iowa State University** *Aug. 2014 - Jun. 2022 (expected)*
  PhD Candidate in Dept. of Computer Science. *Advisor: Prof. Jin Tian*
  Major: Programming Language and System; AI, Machine Learning and NLP
- **University of Science and Technology of China** *Sep. 2010 - Jun. 2014*
  B.E. in Electrical Engineering

Keywords

- Compiler, PL, SE, Embedded DSL, IDE, Fullstack, ReactJS, GraphQL, SQL, Linux, Docker, Kubernetes, Rancher
- Machine Learning, Deep Learning, Neural Network, Tensorflow, Pytorch, NLP, Auto-encoder, Adversarial Learning

Professional Skills

- **Programming**: Python, Julia, Javascript, Racket/Scheme/Lisp, Meta Prog, DSL, C/C++, LLVM, Tmux, Jupyter
- **Fullstack Web**: ReactJS, ExpressJS, GraphQL, SQL, NoSQL, Linux, Container, Docker, K8s, Rancher, Proxmox
- **Machine Learning**: Tensorflow, Pytorch, Flux, MLP, CNN, LSTM, Auto-encoders, Scikit-learn, Numpy, Pandas

Projects & Experience

- **Programming Language (PL) & Software Engineering (SE)**

  [PL, AI] An Embedded Programming Language and System for PCB Designs *Feb. 2020 - Oct. 2021*
  - Designed and implemented **BHDL** [1], the first open-source embedded domain-specific language (eDSL) for creating Printed Circuit Boards (PCBs). BHDL enables engineers to program complex and reusable PCBs.
  - Co-Designed **DRL-MCTS** [2], 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.

  [PL, SE] A Hierarchical IDE for Interactive Development at Scale *Aug. 2020 - Oct. 2021*
  - Designed and developed **CodePod**[3], 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.

- **AI & Machine Learning (ML) & Natural Language Processing (NLP)**

  [AI, ML] Self-supervised DAG Structure Learning *Nov. 2019 - Aug. 2021*
  - Proposed **DAG-EQ** [4], the first self-supervised learning formulation of the DAG structure learning problem.
  - Applied equivariant neural networks to solve the problem and achieved state-of-the-art on DAG learning.
  - Applied transfer learning and ensemble learning and show significantly better scalability to large graphs.

  [NLP] End-to-end Semantics-based Summary Quality Assessment *Feb. 2019 - Feb. 2020*
  - Co-designed a reference-free supervised approach **FreeRouge** [5] 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.

  [ML, DL] Adversarial Attacks and Defense *Aug. 2018 - Sep. 2019*
  - Proposed a novel adversarial auto-encoder **AdvAE** [6] that protects Neural Networks from adversarial attacks.
  - Showed that AdvAE significantly outperforms other purifying-based adversarial defense techniques.

Publications

[1] **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 website: `https://bhdl.org`
[2] Youbiao He, **Hebi Li**, Forrest Bao, *"Circuit Routing Using Monte Carlo Tree Search and Deep Reinforcement Learning"*, arXiv preprint arXiv:2006.13607 (2020). submitted to ISCAS-22.
[3] **Hebi, Li**, F. Bao, Qi Xiao and Jin Tian, *"CodePod: A Namespace-Aware, Hierarchical Jupyter for Interactive Development at Scale"*, submitted to ICSE-22
[4] **Hebi Li**, Qi Xiao, and Jin Tian. *"Supervised Whole DAG Causal Discovery."*, arXiv:2006.04697 (2020). submitted to JMLR/CLEAR-22 Source Code: `https://github.com/lihebi/DAG-EQ`
[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), submitting to NAACL-22.
[6] **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`
[7] Qi Xiao, **H. Li**, J. Tian and Z. Wang. *"Groupwise Feature Selection for Supervised Learning"*, submitted to ICASSP-22