Ruihang **Lai**

✉ ruihangl@cs.cmu.edu | 🏠 ruihanglai.com | MasterJH5574

Education

Carnegie Mellon University   *Pittsburgh, United States*
PH.D. IN COMPUTER SCIENCE   *Aug. 2022 - Present*

Shanghai Jiao Tong University   *Shanghai, P.R. China*
B.ENG. IN COMPUTER SCIENCE   *Sep. 2018 - Jun. 2022*
- Member of ACM Honors Class, a pilot CS program for top talented students.
- Advised by Prof. Yong Yu.

Research Interests

- **Machine Learning Systems (Systems for Large-Scale Workloads)**
- **Machine Learning Compilation**
- **Systems for Emerging Computation (E.g., Sparse Computing)**

Experiences

OctoML   *Jun. 2022 - Aug. 2022*
RESEARCH INTERN
- Working on the Machine Learning Compilation online course, as a teaching assistant.

Catalyst Research Group, Carnegie Mellon University   *Oct. 2021 - Dec. 2021*
RESEARCH INTERN
- Worked on automatic tensor program optimization and sparse tensor computing.
- Advised by Prof. Tianqi Chen.

Publications

SparseTIR: Composable Abstractions for Sparse Compilation in Deep Learning [arxiv]   *In submission. Under review.*

Zihao Ye, *Ruihang Lai*, Junru Shao, Tianqi Chen, Luis Ceze

TensorIR: An Abstraction for Automatic Tensorized Program Optimization [arxiv]   *In submission. Under review.*
Siyuan Feng, Bohan Hou, Hongyi Jin, Wuwei Lin, Junru Shao, *Ruihang Lai*, Zihao Ye, Lianmin Zheng, Cody Hao Yu, Yong Yu, Tianqi Chen

Tensor Program Optimization with Probabilistic Programs [arxiv]   *Accepted by NeurIPS 2022.*
Junru Shao, Xiyou Zhou, Siyuan Feng, Bohan Hou, *Ruihang Lai*, Hongyi Jin, Wuwei Lin, Masahiro Masuda, Cody Hao Yu, Tianqi Chen

Talks

TensorIR: An Abstraction for Tensorized Program Optimization [video]
- Jan. 2022 @ System, Architecture, Machine learning, and Programming language (SAMPL) Lab.

SparseTIR: A Unified Abstraction for Sparse Workload Representation and Optimization
- Nov. 2021 @ CMU Automated Learning System (Catalyst) Group.
- Dec. 2021 @ TVM Conference 2021.

Selected Projects

Apache TVM, An End-to-End Machine Learning Compiler Framework

COMMITTER   *Fall 2020 - Present*

- Author of over 10+ PRs, 6500+ lines of code.
- Reviewer of over 25+ PRs.

Mx-Compiler

COURSE PROJECT   *Spring 2020*

- A toy compiler implemented in Java, from Mx* (a C- and Java-like language) to RISC-V assembly code.
- Implemented many effective optimizations. The generated code has performance close to GCC O2.
- More than 15k+ lines of code overall.

Distributed Hash Table

COURSE PROJECT   *Summer 2019*

- Implemented two DHT protocols, Chord and Kademlia, in Go Language.
- Implemented an instant chat room system based on the Chord protocol.

Teaching

Machine Learning Compilation   *Online course*

TEACHING ASSISTANT   *Summer 2022*

- Prepare and release course assignments. Answer questions in the discussion page.

Principle and Practice of Computer Algorithms   *SJTU*

LEADER TEACHING ASSISTANT   *Summer 2020*

- Advised students to implement a RISC-V simulator.
- Advised students to implement two Distributed Hash Table protocols, Chord and Kademlia, in Go Language.

Data Structure (Honor)   *SJTU*

LEADER TEACHING ASSISTANT   *Spring 2020*

- Taught advanced data structures which students usually do not learn in class.
- Prepared the course assignments, projects and programming exams.

Honors & Awards

SCHOLARSHIP

2022	**Shanghai Excellent Graduate Award**	*Shanghai, China*
2020	**National Scholarship** (Top 0.2% nationwide)	*P.R. China*
2020	**Tang Lixin Scholarship** (Less than 30 undergraduates school-wide each year)	*SJTU*

PROGRAMMING COMPETITIONS

2018	**The 4th Place** The 2018 ICPC Asia Singapore Regional Contests	*Singapore*
2017	**Silver Prize** The 34th China National Olympiad in Informatics (NOI)	*P.R. China*