

LI XIANG

✉ lixiang99410@gmail.com · ☎ (+86) 18652936169

🎓 EDUCATION

Institute of Compute Technology, Chinese Academy of Sciences, Beijing Sept. 2020 – Jun. 2023
Master in Electronic Information (Direction: Computer Architecture) (expected)
Nanjing University of Aeronautics and Astronautics, Nanjing Sept. 2016 – Jun. 2020
B.S. in Computer Science and Technology (Honor Class)

🔖 PROJECTS

Performance tuning of the GPT-2 model with billions of parameters Jul. 2021 – Aug. 2021
Main contributor Mentor: Dr. Yanghua Peng & Bairen Yi from Data-AML in Bytedance

Brief introduction:

- Use DeepSpeed framework to fine-tuning a GPT-2 model with 1.5B parameters on a single machine with 8 GPUs (NVIDIA V100 16GB).
- One of the projects in ByteCamp Summer 2021. (2nd place in final presentation of over 30 projects)

Linearizable concurrent data structure for train ticket sales Oct. 2020 – Dec. 2020
Complete separately Mentor: Associate Professor Yi Lv from SKLCS, IoS, CAS

Brief introduction:

- Final course work of concurrent data structures and multicore programming
- Use atomic variables, CAS primitives, etc. to maintain linearizability, and pass the linearizability test.
- Use bitmap, cache etc. to speedup queries, the final performance is the top 10% of the course.

♡ HONORS AND AWARDS

Finalist - The 2020 ICPC World Final Dec. 2021 (expected)
Outstanding Graduates - College of Computer Science and Technology, NUAA Jun. 2020
Gold Medal - The 2019 ICPC Asia Shenyang Regional Contest Nov. 2019
Gold Medal - The 2019 CCF Collegiate Computer System & Programming Contest (CCF-CCSP) Oct. 2019

👥 EXPERIENCE OF INTERNSHIP

Bytedance Inc. Beijing, China Dec. 2019 – May. 2020
Software Engineer Intern. - Data-System Mentor: Jingtao Jie

Brief introduction:

- Develop serveral APIs of CRUD affairs using Python.
- Build a data flow from data source to Hive through Kafka using Python and HiveSQL.
- Assist to build a data warehouse.

⚙️ SKILLS

- Programming Languages: C == C++ > Python > Java = shell
- Development: Algorithms, Data structures, Git, Docker, SQL, etc.

🔖 MISCELLANEOUS

- GitHub: <https://github.com/SeanLi-OI>
- Blog: <https://sean-li.github.io/>
- Languages: English - Fluent, Mandarin - Native speaker

李想

✉ lixiang99410@gmail.com · ☎ (+86) 18652936169

🎓 教育经历

中国科学院计算技术研究所, 北京, 中国 2020 年 9 月 – 2023 年 6 月
专业硕士 电子信息 (方向: 计算机体系结构) (预计)
南京航空航天大学, 南京, 中国 2016 年 9 月 – 2020 年 6 月
学士 计算机科学与技术 (培优班)

📌 项目

GPT-2 百亿参数超大模型训练探索 2021 年 7 月 – 2021 年 8 月
主要参与 导师: 彭杨华 & 易百忍, 字节跳动-Data-AML 团队

- 基于DeepSpeed 框架在一台 8 卡 (NVIDIA V100 16GB) 物理机上做 150 亿参数 GPT-2 模型的微调。
- 本项目属于 ByteCamp 2021 夏令营高性能计算方向项目, 在终审中获得第二名 (2/30)。

用于列车售票的可线性化并发数据结构 2020 年 10 月 – 2020 年 12 月
独立完成 导师: 吕毅副研究员, 中科院软件所

- 并发数据结构与多核编程课程的期末大作业;
- 使用 Java 中 Atomic 变量, CAS 原语等维护可线性化, 通过可线性化测试;
- 利用位运算, 缓存等加速车票查询, 在 128 核服务器上测试性能在全课程位于前 10%。

♡ 获奖与荣誉

入围 - 2020 年 ICPC 世界总决赛 2021 年 12 月 (预计)
优秀毕业生 - 南京航空航天大学计算机科学与技术学院 2020 年 6 月
金奖 - 2019 年 ICPC 亚洲区域赛沈阳站 2019 年 11 月
金奖 - 2019 年 ICPC 亚洲区域赛银川站 2019 年 10 月
金奖 - 2019 年 CCF 大学生计算机系统与程序设计竞赛 2019 年 10 月
金奖 - 2018 年第一届小米程序设计大赛 2018 年 9 月
全球 108 名 - Google Codejam Kickstart 2018 年 7 月

👨‍💻 实习经历

北京字跳网络技术有限公司 北京, 中国 2019 年 12 月 – 2020 年 5 月
后端研发实习生 Data-系统部 Mentor: 介静涛

- 基于 Python 的若干增删改查的后端 api 开发
- 用 Kafka, HDFS 等大数据基础设施, 通过 HiveSQL 建立了若干稳定的数据清洗工作流。
- 辅助完成在若干数据仓库的搭建工作。

⚙️ 能力

- 编程语言: C == C++ > Python > Java = shell
- 其他能力: 算法, 数据结构, Git, Docker, SQL, 等等。

📌 其他

- GitHub: <https://github.com/SeanLi-OI>
- 博客: <https://sean-li.github.io/>
- 语言: 中文 - 母语, 英文 - 流利。