

孙玉粮

(+86) 139-4510-8536 · 21371245@buaa.edu.cn · <https://intrepidLi.github.io/>

教育背景

北京航空航天大学, 计算机科学与技术, 在读大三学生

2021.9 至今

GPA 3.67/4.0, 87.5/100

技术能力

- 编程语言: Python, C, C++, Java, Verilog, JavaScript, HTML/CSS, SQL
- 工程方面: Git, Flask

项目经历

AirQA 民航旅客问答系统 | 北航 ACT 实验室, 指导人: 张日崇教授

2023.1-2023.4

- 负责构造整套数据处理系统使之能够对航司原始杂乱数据进行分类生成问答对
- 采用模板化方法和 docT5query 模型分别对规律化和非规律化数据进行问答数据生成
- 采用基于百度公司的 RocketQA 预训练模型在民航数据集上训练出的新模型进行检索式问答生成

基于 SysY 的编译器 | 北航编译技术课程设计

2023.9-2023.12

项目地址: https://github.com/intrepidLi/BUAA_Compilers_2023

- 使用 Java 完成编译器, 可对 SysY (C 语言的一个子集) 进行从源码到 LLVM IR 中间代码的转换
 - 系统组成: 词法分析, 语法分析, 语义分析生成抽象语法树 (AST), 中间代码生成
 - 利用自动机理论手动进行词法分析, 利用递归下降进行语法分析生成抽象语法树 (AST), 同时进行错误处理。之后使用“一切皆 Value”的思想构建新的 LLVM IR 的语法树, 生成 LLVM IR 代码
- 本项目是北航编译技术的课程设计, 通过了目标代码为 LLVM IR 编译器的所有测试。

北航比价跳蚤市场 | 北航数据库系统概论课程设计

2023.9-2023.12

项目地址: https://github.com/intrepidLi/BUAA_DataBase_2023

- 使用 ReactJS, Flask 框架和华为云数据库完成的前后端交互网站, 包括用户系统, 评论系统, 商品交易系统等功能
- 主要负责前端和交互工作, 包括使用 Material-UI 进行网站设计, 并用 axios 库进行前后端交互和前后端的 Debug 完善

论文发表

WaterBench: Towards Holistic Evaluation of Watermarks for Large Language Models

Shangqing Tu*, Yuliang Sun*, Yushi Bai, Jifan Yu, Lei Hou, Juanzi Li(* 相同贡献)

- 预印本: <https://arxiv.org/abs/2311.07138>
- ACL2024 在投

研究经历

WaterBench: Towards Holistic Evaluation of Watermarks for Large Language Models

研究助理 | 清华大学知识工程研究室, 指导人: 李涓子教授

2023.5-2023.12

- 为了完整全面地评估大模型生成文本中的水印效果, 我们提出了一种新的思路: 通过超参数搜索将不同水印强度调至相同, 之后测试带有不同水印的文本质量
- 构建了九个不同的任务 (五个类别) 作为评估数据集
- 使用 GPT4-Judge 和人工评价的方法对文本进行评估
- 主要负责超参数搜索, 评估实验的进行, 数据处理以及部分论文的编写

竞赛/荣誉

- 北航 2021-2022 学年学习优秀二等奖学金
- 北航 2021-2022 学年学科竞赛二等奖学金
- 第十四届蓝桥杯 C/C++ 研究生组三等奖, 2023 年 4 月

其他

- 语言: 英语 (CET6), 汉语
- GitHub: <https://github.com/intrepidLi>