CHAO PENG

ByteDance, Dazhongsi Square, Haidian District, Beijing, China
chao.peng@acm.org

EDUCATION

University of Edinburgh **Edinburgh, UK**
PhD in Informatics (Software Engineering) September 2017 - November 2021
Supervised by Dr. Ajitha Rajan

- Fully funded by the School of Informatics Scholarship
- Research fields:
  - Coverage metrics and test case generation for GPU programs,
  - Program analysis of smart contracts,
  - Android GUI testing.

University of Edinburgh **Edinburgh, UK**
MSc in High Performance Computing with Data Science September 2016 - August 2017

- Thesis: FEniCS Performance Investigation and Porting miniDFT to GPU Clusters
- Received Achievement Award of the 6th Annual HPCAC Student Cluster Competition

Xuzhou University of Technology **Xuzhou, China**
BSc in Computer Science and Technology September 2012 - June 2016

- Awarded 2 × National Scholarship by Ministry of Education of China
- Awarded National Encouragement Scholarship by Department of Education of Jiangsu Province
- Awarded Hornored Undergraduate of Xuzhou University of Technology
- President of the Student Union of School of Information & Electrical Engineering

WORK EXPERIENCE

ByteDance 05 Jul 2021 - Present
Software Engineering Lab *Beijing, China*

· Position: Principal Research Scientist
· Duties and Achievements:

- Built and leading a research team conducting cutting-edge software engineering research with publications in top-tier academic conference such as ICSE, FSE, ASE, ACL, etc.
- Data engineering and model training of a reinforcement learning LLM tailored for software engineering tasks, which outperformed other models of the size parameter size.
- Built MarsCode Agent, a general software development and bug fixing LLM-based agent which ranked the first place at the SWE-bench Lite and Verified benchmark.
- Setup collaboration projects with leading universitis such as Tsinghua University, Peking University, Shanghai Jiaotong University, Fudan University, Tianjin University, Chinese University of Hong Kong (Shenzhen), etc.

Fudan University 02 Mar 2022 - Present
School of Computer Science *Shanghai, China*

· Position: Part-time Postgraduate Student Mentor

INTERNSHIP EXPERIENCE

Huawei Edinburgh Research Centre 23/11/2020 - 22/06/2021
Programming Languages Lab, 2021 Laboratories *Edinburgh, UK*

· Mentor: Dan Ghica, Head of PL Lab
· Duties: Compiler backend testing and benchmarking. Compiler implementation and optimisation.

Huawei London Research Centre 28/10/2019 - 28/02/2020
Hisilicon Kirin GPU Team, 2012 Laboratories *London, UK*

- Mentor: Graham Connor, Chief GPU Scientist
- Duties: Design of the layout of the benchmarking framework, implementation of the build system and cross-platform support, design and implementation of GPU micro-benchmarks.

RESEARCH PUBLICATIONS

Pre-print / Under Review

8. Zexiong Ma, **Chao Peng**, Qunhong Zeng, Pengfei Gao, Yanzhen Zou, Bing Xie (2025). *Tool-integrated Reinforcement Learning for Repo Deep Search.* arXiv:2508.03012
7. Xinghang Li, Jingzhe Ding, **Chao Peng**, Bing Zhao, Xiang Gao, Hongwan Gao, Xinchen Gu (2025). *SafeGen-Bench: A Benchmark Framework for Security Vulnerability Detection in LLM-Generated Code.* arXiv:2506.05692
6. Hu, Ruida, **Chao Peng**, Xinchen Wang, and Cuiyun Gao (2025). *An LLM-based Agent for Reliable Docker Environment Configuration.* arXiv:2502.13681
5. Chen, Jialiang, Kaifa Zhao, Jie Liu, **Chao Peng**, Jierui Liu, Hang Zhu, Pengfei Gao, Ping Yang, Shuiguang Deng (2024). *CoReQA: Uncovering Potentials of Language Models in Code Repository Question Answering.* arXiv:2501.03447
4. Guan, Zhanming, Junlin Liu, Jierui Liu, **Chao Peng**, Dexin Liu, Ningyuan Sun, Bo Jiang, Wenchao Li, Jie Liu, and Hang Zhu (2024). *ContextModule: Improving Code Completion via Repository-level Contextual Information.* arXiv:2412.08063
3. Tsimpourlas, Foivos, **Chao Peng**, Carlos Rosuero, Ping Yang, and Ajitha Rajan (2024) *Go-Oracle: Automated Test Oracle for Go Concurrency Bugs.* arXiv:2412.08061
2. Meng, Xiangxin, Zexiong Ma, Pengfei Gao, **Chao Peng** (2024). *An Empirical Study on LLM-based Agents for Automated Bug Fixing.* arXiv:2411.10213
1. Liu, Yizhou, Pengfei Gao, Xinchen Wang, Jie Liu, Yexuan Shi, Zhao Zhang, **Chao Peng** (2024). *MarsCode Agent: AI-native Automated Bug Fixing.* arXiv:2409.00899

Peer-reviewed

25. Wang, Xinchen and Gao, Pengfei and **Peng, Chao** and Hu, Ruida and Gao, Cuiyun (2025). *CodeVisionary: An Agent-based Framework for Evaluating Large Language Models in Code Generation.* In ASE 2025. November 2025.
24. Wu, Qinyun, **Chao Peng**, Pengfei Gao, Ruida Hu, Haoyu Gan, Bo Jiang, Jinhe Tang, Zhiwen Deng, Zhanming Guan, Cuiyun Gao, Xia Liu, Ping Yang (2025). *RepoMasterEval: Evaluating Code Completion via Real-World Repositories.* In ASE 2025 Industry Track. November 2025.
23. Tsimpourlas, Foivos, **Chao Peng**, Carlos Rosuero, Ping Yang, Ajitha Rajan (2025). *Go-Oracle: Automated Test Oracle for Go Concurrency Bugs.* In ESEM 2025. October 2025.
22. Ma, Zexiong, **Chao Peng**, Pengfei Gao, Xiangxin Meng, Yanzhen Zou, and Bing Xie (2025). *SoRFT: Issue Resolving with Subtask-oriented Reinforced Fine-Tuning.* In ACL 2025. July 2025.
21. Ruida Hu, **Chao Peng**, Jingyi Ren, Bo Jiang, Xiangxin Meng, Qinyun Wu, Pengfei Gao, Xinchen Wang and Cuiyun Gao (2025). *Understanding Large Language Model Performance in Software Engineering: A Large-scale Question Answering Benchmark.* In SIGIR 2025 Short Paper Track. July 2025
20. Xinchen Wang, Pengfei Gao, Xiangxin Meng, **Chao Peng**, Ruida Hu, Yun Lin, Cuiyun Gao (2025). *AEGIS: An Agent-based Framework for Bug Reproduction from Issue Descriptions.* In FSE 2025 Industry Track. June 2025
19. Li, Bowen, Wenhan Wu, Ziwei Tang, Lin Shi, John Yang, Jinyang Li, Shunyu Yao, Chen Qian, Binyuan Hui, Qicheng Zhang, Zhiyin Yu, He Du, Ping Yang, Dahua Lin, **Chao Peng**, Kai Chen (2025). *Prompting Large Language Models to Tackle the Full Software Development Lifecycle: A Case Study.* In proceedings of the 31st International Conference on Computational Linguistics (COLING 2025). January 2025
18. Liang, Xiaoyun, Jingyi Ren, Jiayi Qi, **Chao Peng**, and Bo Jiang. *DialogAgent: An Auto-engagement Agent for Code Question Answering Data Production.* In ICSE 2025 SEIP Track April 2024
17. Chao Peng, Qinyun Wu, Jiangchao Liu, Jierui Liu, Bo Jiang, Mengqian Xu, Yinghao Wang, Xia Liu, Ping

Yang (2024). *RepoSim: Evaluating Prompt Strategies for Code Completion via User Behavior Simulation.* In proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering (ASE 2024), New Idea and Emerging Results Track October 2024

16. Zhu Tao, Yongqiang Gao, Jiayi Qi, **Chao Peng**, Qinyun Wu, Xiang Chen, Ping Yang (2024). *Neat: Mobile App Layout Similarity Comparison based on Graph Convolutional Networks.* In proceedings of the ACM International Conference on the Foundations of Software Engineering (FSE 2024), Industry Track July 2024

15. Chao Peng, Zhengwei Lv, Jiarong Fu, Jiayuan Liang, Zhao Zhang, Ajitha Rajan, Ping Yang(2024). *Hawkeye: Change-targeted Testing for Android Apps based on Deep Reinforcement Learning.* In proceedings of the 46th IEEE/ACM International Conference on Software Engineering, Software Engineering in Practice Track (ICSE SEIP 2024) April 2024

14. Xiaoyun Liang, Jiayi Qi, Yongqiang Gao, **Chao Peng**, Ping Yang (2023). *AG3: Automated Game GUI Text Glitch Detection based on Computer Vision.* In proceedings of the 2023 ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE 2023) December 2023

13. Zongze Jiang, Ming Wen, Yixin Yang, **Chao Peng**, Ping Yang, Hai Jin, (2023). *Effective Concurrency Testing for Go via Directional Primitive-constrained Interleaving Exploration.* In proceedings of the 38th IEEE/ACM International Conference on Automated Software Engineering (ASE 2023) September 2023

12. Siwei Wang, Xue Mao, Ziguang Gao, Yujun Gao, Qucheng Shen, **Chao Peng** (2023). *NxtUnit: Automated Unit Test Generation for Go.* In proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering (EASE 2023). June 2023

11. Jingling Sun, Ting Su, Kai Liu, **Chao Peng**, Zhao Zhang, Geguang Pu, Tao Xie, and Zhendong Su (2022). *Characterizing and Finding SystemSetting-Related Defects in Android Apps.* IEEE Transactions on Software Engineering (TSE). April 2023

10. Zhengwei Lv, **Chao Peng**, Zhao Zhang, Ting Su, Kai Liu, Ping Yang. *Fastbot2: Reusable Automated Model-based GUI Testing for Android Enhanced by Reinforcement Learning.* In proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering (ASE 2022) Industry Showcase Track. October 2022

9. Chao Peng, Zhao Zhang, Zhengwei Lv and Ping Yang. *MUBot: Learning to Test Large-Scale Commercial Android Apps like a Human.* In proceedings of the 38th International Conference on Software Maintenance and Evolution (ICSME 2022) Industry Track October 2022

8. Chao Peng, Yujun Gao and Ping Yang. *Automated Server Testing: an Industrial Experience Report.* In proceedings of the 38th International Conference on Software Maintenance and Evolution (ICSME 2022) Industry Track October 2022

7. Sefa Akca, **Chao Peng**, Ajitha Rajan. *Testing Smart Contracts: Which Technique Performs Best?* In proceedings of the 15th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM 2021) October 2021

6. Chao Peng, Ajitha Rajan, Tianqin Cai. *CAT: Change-focused Android GUI Testing.* In proceedings of the 37th International Conference on Software Maintenance and Evolution (ICSME 2021) September 2021

5. Chao Peng, Ajitha Rajan. *Automated Test Generation for OpenCL Kernels using Fuzzing and Constraint Solving.* In proceedings of the 11th Workshop on General Purpose GPUs 2020 (GPGPU @ PPoPP) Feburary 2020

4. Chao Peng, Sefa Akca, Ajitha Rajan. *SIF: A Framework for Solidity Contract Instrumentation and Analysis.* In proceedings of the the 26th Asia-Pacific Software Engineering Conference (APSEC 2019). December 2019

3. Sefa Akca, Ajitha Rajan, **Chao Peng**. *SolAnalyser: A Framework for Analysing and Testing Smart Contracts.* In proceedings of the the 26th Asia-Pacific Software Engineering Conference (APSEC 2019). December 2019

2. Chao Peng. *On the Correctness of GPU Programs.* In proceedings of the 2019 ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2019). July 2019

1. Chao Peng, Ajitha Rajan. *CLTestCheck: Measuring Test Effectiveness for GPU Kernels.* In proceedings of the 22nd International Conference on Fundamental Approaches to Software Engineering (FASE 2019). April 2019

PROFESSIONAL SERVICE

Organising/Steering Committee	AI-IDE@FSE'25, APR@ICSE'25, AIWare
PC Member (Research/Technical Track)	LLM4Code'26, FSE'26, ASE'25, LLM4Code'25, FSE'25, MSR'23(Junior PC), A-Mobile'22, PRDC'22
PC Member (Industry Track)	SANER'26, ASE'25, ISSRE'24, APSEC'24, MSR'24
PC Member (Artifact Evaluation)	OSDI'22, USENIX ATC'22, ISSTA'22, 21, 20
Student Volunteer	PLDI'20, ISSTA'19, ETAPS'19
Reviewer	ICSE'19 SEIP (Software Engineering in Practice) Track

SKILLS

Languages	Mandarin Chinese - Native English - Proficient German - Basic
Technical Skills	Programming languages: C/C++, Java, Fortran, Python, Kotlin, Rust, Go Parallel programming: OpenMP, MPI, OpenCL, CUDA, SYCL Program Analysis: Java Soot, Clang LibTooling Computer Graphics: OpenGL (ES), Vulkan
Tools	LaTeX, Git, SVN, CMake, Gradle

TEACHING EXPERIENCE

MSc Projects Apr 18 - Aug 18 and Apr 19 - Aug 19
Sub-supervisor for MSc Projects *School of Informatics, University of Edinburgh*

Helping supervise the following MSc students for their projects:
- Bowen Du: Code coverage measurement for GPU programs
- Yangning Li: Random test case generation for OpenCL programs

Software Testing Jan 18 - Apr 18, Jan 19 - Apr 19 and Jan 21 - Apr 21
Teaching Assistant, Tutor and Marker *School of Informatics, University of Edinburgh*

Computer Programming Skills and Concepts Nov 17 - Dec 17 and Dec 18 - Jan 19
Marker *School of Informatics, University of Edinburgh*

Introduction to Java Programming Sep 17 - Dec 17 and Sep 18 - Dec 18
Demonstrator *School of Informatics, University of Edinburgh*

Nominated for Teaching Awards by students of this course

SCHOLARSHIPS & HONOURS

Scholarships

- School of Informatics Scholarship (fully-funded PhD scholarship) Apr 17
- Outstanding Undergraduate Scholarship Jun 16
- China National Scholarship of the 2014/15 Academic Year Nov 15
- China National Encouragement Scholarship of the 2013/14 Academic Year Nov 14
- China National Scholarship of the 2012/13 Academic Year Nov 13

Awards

- Nomination for Teaching Awards: Best Student Who Tutors Award Feb 18
- Certificate of Achievement for Participating in the 6th Annual HPCAC Student Cluster Competition Jun 17
- Distinct Graduate of Xuzhou University of Technology Jun 16
- Top Ten Role Models of Learning at Xuzhou University of Technology Mar 16
- Jiangsu Province-level Excellent Student Cadre May 15