2025 IEEE/ACM 33rd International Conference on Program Comprehension (ICPC 2025)

Ottawa, Ontario, Canada
27-28 April 2025

IEEE

IEEE Catalog Number: CFP25009-POD
ISBN: 979-8-3315-0224-9

Copyright © 2025 by the Institute of Electrical and Electronics Engineers, Inc.
All Rights Reserved

Copyright and Reprint Permissions: Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923.

For other copying, reprint or republication permission, write to IEEE Copyrights Manager, IEEE Service Center, 445 Hoes Lane, Piscataway, NJ 08854. All rights reserved.

****** This is a print representation of what appears in the IEEE Digital Library. Some format issues inherent in the e-media version may also appear in this print version.***

IEEE Catalog Number:	CFP25009-POD
ISBN (Print-On-Demand):	979-8-3315-0224-9
ISBN (Online):	979-8-3315-0223-2
ISSN:	2643-7147

Additional Copies of This Publication Are Available From:

Curran Associates, Inc
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: (845) 758-0400
Fax: (845) 758-2633
E-mail: curran@proceedings.com
Web: www.proceedings.com

2025 IEEE/ACM 33rd International Conference on Program Comprehension (ICPC)

ICPC 2025

Table of Contents

Message from the ICPC 2025 Chairs ..... **xiii**
Organizing Committee ..... **xv**
Steering Committee ..... **xvi**
Program Committees ..... **xvii**
Keynotes ..... **xxi**

App Comprehension

Combining Language and App UI Analysis for the Automated Assessment of Bug Reproduction Steps 1
Junayed Mahmud (University of Central Florida (USA)), Antu Saha (William & Mary (USA)), Oscar Chaparro (William & Mary (USA)), Kevin Moran (University of Central Florida (USA)), and Andrian Marcus (George Mason University (USA))

Effectively Modeling UI Transition Graphs for Android Apps via Reinforcement Learning 13
Wunan Guo (University of Shanghai for Science and Technology, China), Zhen Dong (Fudan University, China), Liwei Shen (Fudan University, China), Daihong Zhou (Shanghai Institute of Technology, China), Bin Hu (Hangzhou Dianzi University, China), Chen Zhang (Soochow University, China), and Hai Xue (University of Shanghai for Science and Technology, China)

Characterizing Bugs in Login Processes of Android Applications: An Empirical Study 25
Zixu Zhou (McGill University, Canada), Rufeng Chen (McGill University, Canada), Junfeng Chen (Southern University of Science and Technology, China), Yepang Liu (Southern University of Science and Technology, China), and Lili Wei (McGill University, Canada)

Vulnerabilities, Technical Debt, Defects

CalmDroid: Core-Set Based Active Learning for Multi-Label Android Malware Detection 37
Minhong Dong (Tiangong University, China), Liyuan Liu (Tiangong University, China), Mengting Zhang (Tiangong University, China), Sen Chen (Nankai University, China), Wenying He (Hebei University of Technology, China), Ze Wang (Tiangong University, China), and Yude Bai (Tiangong University, China)

Towards Task-Harmonious Vulnerability Assessment Based on LLM 49
Zaixing Zhang (Southeast University, China), Jianming Chang (Southeast University, China), Tianyuan Hu (Southeast University, China), Lulu Wang (Southeast University, China), and Bixin Li (Southeast University, China)

A Slicing-Based Approach for Detecting and Patching Vulnerable Code Clones 60
Hakam W. Alomari (Miami University, USA), Christopher Vendome (Miami University, USA), and Himal Gyawali (Miami University, USA)

Revisiting Security Practices for GitHub Actions Workflows 73
Jiangnan Huang (Radboud University, The Netherlands) and Bin Lin (Hangzhou Dianzi University, China)

Leveraging Multi-Task Learning to Improve the Detection of SATD and Vulnerability 78
Barbara Russo (Free University of Bozen-Bolzano, Italy), Jorge Melegati (Free University of Bozen-Bolzano, Italy), and Moritz Mock (Free University of Bozen-Bolzano, Italy)

Leveraging Context Information for Self-Admitted Technical Debt Detection 90
Miki Yonekura (Nara Institute of Science and Technology, Japan), Yutaro Kashiwa (Nara Institute of Science and Technology, Japan), Bin Lin (Hangzhou Dianzi University, China), Kenji Fujiwara (Nara Women's University, Japan), and Hajimu Iida (Nara Institute of Science and Technology, Japan)

Personalized Code Readability Assessment: Are We There Yet? 102
Antonio Vitale (DEVISER @ University of Molise, Italy; Politecnico di Torino, Italy), Emanuela Guglielmi (DEVISER @ University of Molise, Italy), Rocco Oliveto (DEVISER @ University of Molise, Italy), and Simone Scalabrino (DEVISER @ University of Molise, Italy)

Automated Refactoring of Non-Idiomatic Python Code: A Differentiated Replication with LLMs... 113
Alessandro Midolo (University of Sannio, Italy) and Massimiliano Di Penta (University of Sannio, Italy)

SemBug: Detecting Logic Bugs in DBMS through Generating Semantic-Aware Non-Optimizing Query 124
Shiyang Ye (Zhejiang University, China), Chao Ni (Zhejiang University, China), Jue Wang (ETH Zurich, Switzerland), Qianqian Pang (Zhejiang University, China), Xinrui Li (Zhejiang University, China), and Xiaodan Xu (Zhejiang University, China)

A Study on Applying Large Language Models to Issue Classification 136
Jueun Heo (Gyeongsang National University, Republic of Korea) and Seonah Lee (Gyeongsang National University, Republic of Korea)

Education, Debugging, Dynamic Analysis

JavaWiz: A Trace-Based Graphical Debugger for Software Development Education 147
Markus Weninger (Johannes Kepler University, Austria), Simon Grünbacher (Johannes Kepler University, Austria), and Herbert Prähofer (Johannes Kepler University, Austria)

Pinpointing the Learning Obstacles of an Interactive Theorem Prover 159
Sára Juhošová (Delft University of Technology, The Netherlands), Andy Zaidman (Delft University of Technology, The Netherlands), and Jesper Cockx (Delft University of Technology, The Netherlands)

AI-Based Automated Grading of Source Code of Introductory Programming Assignments 171
Jayant Havare (IIT Bombay, India), Varsha Apte (IIT Bombay, India), Kaushikraj Maharajan (IIT Bombay, India), Nithin Chandra Gupta Samudrala (IIT Bombay, India), Ganesh Ramakrishnan (IIT Bombay, India), Srikanth Tamilselvam (IBM Research, India), and Sainath Vavilapalli (IIT Bombay, India)

Students' Program Comprehension Processes in a Large Code Base 182
Anshul Shah (University of California, San Diego, CA), Thanh Tong (University of California, San Diego, CA), Elena Tomson (University of California, San Diego, CA), Steven Shi (University of California, San Diego, CA), William G. Griswold (University of California, San Diego, CA), and Adalbert Gerald Soosai Raj (University of California, San Diego, CA)

OVERLORD: A C++ Overloading Inspector 194
Botond István Horváth (ELTE Eötvös Loránd University, Hungary), Richárd Szalay (ELTE Eötvös Loránd University, Hungary), and Zoltán Porkoláb (ELTE Eötvös Loránd University, Hungary)

Optimizing Code Runtime Performance through Context-Aware Retrieval-Augmented Generation 199
Manish Acharya (Vanderbilt University, USA), Yifan Zhang (Vanderbilt University, USA), Kevin Leach (Vanderbilt University, USA), and Yu Huang (Vanderbilt University, USA)

Investigating Execution-Aware Language Models for Code Optimization 204
Federico Di Menna (University of L'Aquila, Italy), Luca Traini (University of L'Aquila, Italy), Gabriele Bavota (Università della Svizzera italiana, Switzerland), and Vittorio Cortellessa (University of L'Aquila, Italy)

Understanding Data Access in Microservices Applications using Interactive Treemaps 216
Maxime André (University of Namur, Belgium), Marco Raglianti (USI, Switzerland), Anthony Cleve (University of Namur, Belgium), and Michele Lanza (USI, Switzerland)

Divergence-Driven Debugging: Understanding Behavioral Changes Between Two Program Versions.. 221
Rémi Dufloer (Univ. Lille, France), Imen Sayar (Univ. Lille, France), Anne Etien (Univ. Lille, France), and Steven Costiou (Univ. Lille, France)

KotSuite: Unit Test Generation for Kotlin Programs in Android Applications 226
Feng Yang (Wuhan University, China), Qi Xin (Wuhan University, China), Zhilei Ren (Dalian University of Technology, China), and Jifeng Xuan (Wuhan University, China)

Summarisation, Natural Language Generation

Optimizing Datasets for Code Summarization: Is Code-Comment Coherence Enough? 237
Antonio Vitale (Politecnico di Torino, Italy; University of Molise, Italy), Antonio Mastropaolo (William & Mary, USA), Rocco Oliveto (University of Molise, Italy), Massimiliano Di Penta (University of Sannio, Italy), and Simone Scalabrino (University of Sannio, Italy)

CMDeSum: A Cross-Modal Deliberation Network for Code Summarization 250
Zhifang Liao (Central South University, China), Xiaoyu Liu (Central South University, China), Peng Lan (Central South University, China), Song Yu (Central South University, China), and Pei Liu (Central South University, China)

CLCoSum: Curriculum Learning-Based Code Summarization for Code Language Models 262
Hongkui He (South China University of Technology), Jiexin Wang (South China University of Technology), Liuwen Cao (South China University of Technology), and Yi Cai (MOE of China)

DLCoG: A Novel Framework for Dual-Level Code Comment Generation Based on Semantic Segmentation and In-Context Learning 275
Zhiyang Zhang (Central South University, China), Haiyang Yang (Central South University, China), Qingyang Yan (Central South University, China), Hao Yan (Central South University, China), Weihuan Min (Central South University, China), Zhao Wei (Tencent, China), Li Kuang (Central South University, China), and Yingjie Xia (Hangzhou Dianzi University, China)

Explaining GitHub Actions Failures with Large Language Models: Challenges, Insights, and Limitations 286
Pablo Valenzuela-Toledo (University of Bern, Switzerland; Universidad de La Frontera, Chile), Chuyue Wu (University of Bern, Switzerland), Sandro Hernández (University of Bern, Switzerland), Alexander Boll (University of Bern, Switzerland), Roman Machacek (University of Bern, Switzerland), Sebastiano Panichella (University of Bern, Switzerland), and Timo Kehrer (University of Bern, Switzerland)

Large Language Models are Qualified Benchmark Builders: Rebuilding Pre-Training Datasets for Advancing Code Intelligence Tasks 298
Kang Yang (National University of Defense Technology, China), Xinjun Mao (National University of Defense Technology, China), Shangwen Wang (National University of Defense Technology, China), Yanlin Wang (Sun Yat-sen University), Tanghaoran Zhang (National University of Defense Technology, China), Bo Lin (National University of Defense Technology, China), Yihao Qin (National University of Defense Technology, China), Zhang Zhang (National University of Defense Technology, China), Yao Lu (National University of Defense Technology, China), and Kamal Al-Sabahi (College of Banking and Financial Studies, Oman)

Extracting Formal Specifications from Documents using LLMs for Test Automation 310
Hui Li (Fudan University, China), Zhen Dong (Fudan University, China), Siao Wang (Fudan University, China), Hui Zhang (Fudan University, China), Liwei Shen (Fudan University, China), Xin Peng (Fudan University, China), and Dongdong She (The Hong Kong University of Science and Technology, China)

Using Large Language Models to Generate Concise and Understandable Test Case Summaries 322
Natanael Djajadi (Delft University of Technology, The Netherlands), Amirhossein Deljouyi (Delft University of Technology, The Netherlands), and Andy Zaidman (Delft University of Technology, The Netherlands)

Towards Generating the Rationale for Code Changes 327
Francesco Casillo (University of Salerno), Antonio Mastropaolo (William & Mary), Gabriele Bavota (Universita della Svizzera Italiana), Vincenzo Deufemia (University of Salerno), and Carmine Gravino (University of Salerno)

Empirical Findings, Future Visions, Recommendations

Terminal Lucidity: Envisioning the Future of the Terminal 339
Michael MacInnis (Carleton University, Canada), Olga Baysal (Carleton University, Canada), and Michele Lanza (USI, Switzerland)

Exploring Code Comprehension in Scientific Programming: Preliminary Insights from Research Scientists 350
Alyssia Chen (University of Hawai'i at Mānoa, USA), Carol Wong (University of Hawai'i at Mānoa, USA), Bonita Sharif (University of Nebraska - Lincoln, USA), and Anthony Peruma (University of Hawai'i at Mānoa, USA)

Method Names in Jupyter Notebooks: An Exploratory Study 355
Carol Wong (University of Hawai'i at Mānoa, USA), Gunnar Larsen (University of Hawai'i at Mānoa, USA), Rocky Huang (University of Hawai'i at Mānoa, USA), Bonita Sharif (University of Nebraska - Lincoln, USA), and Anthony Peruma (University of Hawai'i at Mānoa, USA)

SCALAR: A Part-of-Speech Tagger for Identifiers 367
Christian D. Newman (Rochester Institute of Technology, USA), Brandon Scholten (Kent State University, USA), Sophia Testa (Kent State University, USA), Joshua A. C. Behler (Kent State University, USA), Syreen Banabilah (Kent State University, USA), Michael L. Collard (University of Akron, USA), Michael J. Decker (Bowling Green State University, USA), Mohamed Wiem Mkaouer (University of Michigan Flint, USA), Marcos Zampieri (George Mason University, USA), Eman Abdullah AlOmar (Stevens Institute of Technology, USA), Reem Alsuhaibani (Prince Sultan University, Saudi Arabia), Anthony Peruma (University of Hawaii at Manoa, USA), and Jonathan I. Maletic (Kent State University, USA)

How do Papers Make into Machine Learning Frameworks: A Preliminary Study on TensorFlow ... 372
Federica Pepe (University of Sannio, Italy), Claudia Farkas (York University, Canada), Maleknaz Nayebi (York University, Canada), Giuliano Antoniol (Polytechnique Montréal, Canada), and Massimiliano Di Penta (University of Sannio, Italy)

Toward Neurosymbolic Program Comprehension 377
Alejandro Velasco (William & Mary), Aya Garryyeva (William & Mary), David N. Palacio (William & Mary), Antonio Mastropaolo (William & Mary), and Denys Poshyvanyk (William & Mary)

Combining Static Analysis Techniques for Program Comprehension using Slicito 382
Robert Husák (Charles University, Czech Republic), Jan Kofroň (Charles University, Czech Republic), and Filip Zavoral (Charles University, Czech Republic)

Mining Code Change Patterns in Ada Projects 387
Robin van Straeten (Radboud University, The Netherlands) and Bin Lin (Hangzhou Dianzi University, China)

Telling Software Evolution Stories with Sonification 398
Carmen Armenti (USI, Switzerland) and Michele Lanza (USI, Switzerland)

Attributed Multiplex Learning for Analogical Third-Party Library Recommendation and Retrieval 403
Baihui Sang (Nanjing University, China), Liang Wang (Nanjing University, China), Jierui Zhang (Nanjing University, China), and Xianping Tao (Nanjing University, China)

LLM2FedLLM - A Tool for Simulating Federated LLMs for Software Engineering Tasks 414
Jahnavi Kumar (Indian Institute of Technology Tirupati, India), Siddhartha Gandu (Indian Institute of Technology Tirupati, India), and Sridhar Chimalakonda (Indian Institute of Technology Tirupati, India)

Code Generation

Code Ranking with Structure Awareness Contrastive Learning 419
Hailin Huang (South China University of Technology; MOE of China), Liuwen Cao (South China University of Technology; MOE of China), Jiexin Wang (South China University of Technology; MOE of China), Tianchen Yu (South China University of Technology; MOE of China), and Yi Cai (MOE of China; South China University of Technology)

Algorithmic Inversion: A Learnable Algorithm Representation for Code Generation 431
Zhongyi Shi (Chinese Academy of Sciences, China), Fuzhang Wu (Chinese Academy of Sciences, China), Weibin Zeng (Chinese Academy of Sciences, China), Yan Kong (Chinese Academy of Sciences, China), Sicheng Shen (Chinese Academy of Sciences, China), and Yanjun Wu (Chinese Academy of Sciences, China)

Studying How Configurations Impact Code Generation in LLMs: the Case of ChatGPT 442
Benedetta Donato (University of Milano-Bicocca, Italy), Leonardo Mariani (University of Milano-Bicocca, Italy), Daniela Micucci (University of Milano-Bicocca, Italy), and Oliviero Riganelli (University of Milano-Bicocca, Italy)

Quality In, Quality Out: Investigating Training Data's Role in AI Code Generation 454
Cristina Improta (University of Naples Federico II, Italy), Rosalia Tufano (Università della Svizzera italiana (USI), Switzerland), Pietro Liguori (University of Naples Federico II, Italy), Domenico Cotroneo (University of Naples Federico II, Italy), and Gabriele Bavota (Università della Svizzera italiana (USI), Switzerland)

Advancing Large Language Models in Code Generation: USACO Benchmark and Bug Mitigation Insights 466
Jacob Trentini (Monte Vista High School), Victor Liu (Seven Lakes High School), Yiming Peng (Vandegrift High School), and Ziliang Zong (Texas State University)

Enhancing Code Generation for Low-Resource Languages: No Silver Bullet 478
Alessandro Giagnorio (Università della Svizzera italiana, Switzerland), Alberto Martin-Lopez (Università della Svizzera italiana, Switzerland), and Gabriele Bavota (Università della Svizzera italiana, Switzerland)

COFT: Making Large Language Models Better zero-shot Learners for Code Generation 489
Weijia Li (Institute of Software Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; Nanjing Institute of Software Technology, China), Yongjie Qian (North China Electric Power University, China), Ke Gao (Institute of Software Chinese Academy of Sciences, China), Haixin Chen (Institute of Computing Technology, Chinese Academy of Sciences, China), Xinyu Wang (Institute of Software Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China; Nanjing Institute of Software Technology, China), Yuchen Tong (Institute of Computing Technology, Chinese Academy of Sciences, China), Ling Li (Institute of Software Chinese Academy of Sciences, China; University of Chinese Academy of Sciences, China), Yanjun Wu (Institute of Software Chinese Academy of Sciences, China), and Chen Zhao (Institute of Software Chinese Academy of Sciences, China)

On the Possibility of Breaking Copyleft Licenses When Reusing Code Generated by ChatGPT 500
Gaia Colombo (University of Milano-Bicocca, Italy), Leonardo Mariani (University of Milano-Bicocca, Italy), Daniela Micucci (University of Milano-Bicocca, Italy), and Oliviero Riganelli (University of Milano-Bicocca, Italy)

Log Parsing, Bug Localisation, Review Comprehension

Comprehensive Taxonomy for Advanced Log Parsing 512
Issam Sedki (Concordia University, Canada), Abdelwahab Hamou-Lhadj (Concordia University, Canada), Otmane Ait-Mohamed (Concordia University, Canada), and Naser Ezzati-Jivan (Brock University, Canada)

GELog: A GPT-Enhanced Log Representation Method for Anomaly Detection 524
Wenwu Xu (University of Chinese Academy of Sciences, China), Peng Wang (Chinese Academy of Sciences, China), Haichao Shi (Chinese Academy of Sciences, China), Guoqiao Zhou (Chinese Academy of Sciences, China), Junliang Yao (Chinese Academy of Sciences, China), and Xiao-Yu Zhang (Chinese Academy of Sciences, China)

Log Parsing using LLMs with Self-Generated In-Context Learning and Self-Correction 536
Yifan Wu (Peking University, China), Siyu Yu (Peking University, China), and Ying Li (Peking University, China)

LLM-BL: Large Language Models are Zero-Shot Rankers for Bug Localization 548
Zhengliang Li (Nanjing University, China), Zhiwei Jiang (Nanjing University, China), Qiguo Huang (NanJing Audit University, China), and Qing Gu (Nanjing University, China)

Improved IR-Based Bug Localization with Intelligent Relevance Feedback 560
Asif Mohammed Samir (Dalhousie University) and Mohammad Masudur Rahman (Dalhousie University)

Towards Enhancing IR-Based Bug Localization Leveraging Texts and Multimedia from Bug Reports 572
Shamima Yeasmin (University of Saskatchewan), Chanchal K. Roy (University of Saskatchewan), Kevin Schneider (University of Saskatchewan), Mohammad Masudur Rahman (Dalhousie University), Kartik Mittal (University of Saskatchewan), and Ryder Hardy (University of Saskatchewan)

Building Bridges, Not Walls: Fairness-Aware and Accurate Recommendation of Code Reviewers via LLM-Based Agents Collaboration 577
Luqiao Wang (Xidian University, China), Qingshan Li (Xidian University, China), Di Cui (Xidian University, China), Mingkang Wang (Xidian University, China), Yutong Zhao (University of Central Missouri, United States), Yongye Xu (Xidian University, China), Huiying Zhuang (Xidian University, China), Yangtao Zhou (Xidian University, China), and Lu Wang (Xidian University, China)

Code Review Comprehension: Reviewing Strategies Seen Through Code Comprehension Theories 589
Pavlína Wurzel Gonçalves (University of Zurich, Switzerland), Pooja Rani (University of Zurich, Switzerland), Margaret-Anne Storey (University of Victoria, Canada), Diomidis Spinellis (Athens University of Economics and Business, Greece), and Alberto Bacchelli (University of Zurich, Switzerland)

Author Index **603**