

Table 1. List of Included Primary Studies

ID	Year	Title	Venue	Reference
1	2020	Evaluating representation learning of code changes for predicting patch correctness in program repair	ASE	[127]
2	2021	Applying CodeBERT for Automated Program Repair of Java Simple Bugs	MSR	[93]
3	2021	CURE Code-Aware Neural Machine Translation for Automatic Program Repair	ICSE	[56]
4	2021	Generating Bug-Fixes Using Pretrained Transformers	PLDI	[19]
5	2021	Is this Change the Answer to that Problem? Correlating Descriptions of Bug and Code Changes for Evaluating Patch Correctness	ASE	[130]
6	2021	TFix-Learning to Fix Coding Errors with a Text-to-Text Transformer	PMLR	[7]
7	2022	An Analysis of the Automatic Bug Fixing Performance of ChatGPT	APR	[123]
8	2022	An empirical study of deep transfer learning-based program repair for Kotlin projects	FSE/ESEC	[63]
9	2022	Can OpenAI’s codex fix bugs?: an evaluation on QuixBugs	APR	[113]
10	2022	CIRCLE: Continual Repair across Programming Languages	ISSTA	[167]
11	2022	DEAR A Novel Deep Learning-based Approach for Automated Program Repair	ICSE	[76]
12	2022	DeepDev-PERF: A Deep Learning-Based Approach for Improving Software Performance	FSE/ESEC	[33]
13	2022	Fix Bugs with Transformer through a Neural-Symbolic Edit Grammar	ICLR	[48]
14	2022	Framing Program Repair as Code Completion	APR	[115]
15	2022	Less training, more repairing please: revisiting automated program repair via zero-shot learning	FSE/ESEC	[150]
16	2022	Patch Generation with Language Models: Feasibility and Scaling Behavior	ICLR	[64]
17	2022	SYNSHINE: improved fixing of Syntax Errors	TSE	[2]
18	2022	Towards JavaScript program repair with Generative Pre-trained Transformer (GPT-2)	APR	[70]
19	2022	VulRepair: A T5-Based Automated Software Vulnerability Repair	FSE/ESEC	[29]
20	2023	A Chain of AI-based Solutions for Resolving FQNs and Fixing Syntax Errors in Partial Code	arxiv	[53]
21	2023	A critical review of large language model on software engineering: An example from chatgpt and automated program repair	arxiv	[176]
22	2023	Agentcoder: Multi-agent-based code generation with iterative testing and optimisation	arxiv	[49]
23	2023	An Empirical Study of Adoption of ChatGPT for Bug Fixing among Professional Developers	ITA	[35]
24	2023	An Empirical Study on Fine-Tuning Large Language Models of Code for Automated Program Repair	ASE	[50]
25	2023	An Evaluation of the Effectiveness of OpenAI’s ChatGPT for Automated Python Program Bug Fixing using QuixBugs	iSEMANTIC	[146]
26	2023	An Extensive Study on Model Architecture and Program Representation in the Domain of Learning-based Automated Program Repair	APR	[44]
27	2023	Automated Code Editing with Search-Generate-Modify	TSE	[80]
28	2023	Automated Program Repair in the Era of Large Pre-trained Language Models	ICSE	[149]
29	2023	Automated Program Repair Using Generative Models for Code Infilling	AIED	[68]
30	2023	Automated Repair of Programs from Large Language Models	ICSE	[24]
31	2023	Baldur: Whole-Proof Generation and Repair with Large Language Models	FSE/ESEC	[27]
32	2023	Can LLMs Patch Security Issues	arxiv	[4]
33	2023	Coffee: Boost Your Code LLMs by Fixing Bugs with Feedback	arxiv	[96]

Continued on next page

ID	Year	Title	Venue	Reference
34	2023	Copiloting the Copilots: Fusing Large Language Models with Completion Engines for Automated Program Repair	FSE/ESEC	[141]
35	2023	Enhancing Automated Program Repair through Fine-tuning and Prompt Engineering	arxiv	[110]
36	2023	Enhancing Code Language Models for Program Repair by Curricular Fine-tuning Framework	ICSME	[41]
37	2023	Evaluating ChatGPT for Smart Contracts Vulnerability Correction	COMPSAC	[98]
38	2023	Evaluating Pre-trained Language Models for Repairing API Misuses	arxiv	[177]
39	2023	Examining zero-shot vulnerability repair with large language models	S&P	[111]
40	2023	Exploring the Limits of ChatGPT in Software Security Applications	arxiv	[143]
41	2023	FixEval: Execution-based Evaluation of Program Fixes for Programming Problems	APR	[42]
42	2023	Fully Autonomous Programming with Large Language Models	GECCO	[87]
43	2023	Gamma: Revisiting Template-Based Automated Program Repair Via Mask Prediction	ASE	[175]
44	2023	GPT-3-Powered Type Error Debugging: Investigating the Use of Large Language Models for Code Repair	SLE	[116]
45	2023	How Effective Are Neural Networks for Fixing Security Vulnerabilities	ISSTA	[144]
46	2023	Impact of Code Language Models on Automated Program Repair	ICSE	[55]
47	2023	Improving Automated Program Repair with Domain Adaptation	TOSEM	[189]
48	2023	InferFix: End-to-End Program Repair with LLMs over Retrieval-Augmented Prompts	FSE/ESEC	[59]
49	2023	Invalidator: Automated Patch Correctness Assessment via Semantic and Syntactic Reasoning	TSE	[72]
50	2023	Is ChatGPT the Ultimate Programming Assistant-How far is it?	arxiv	[129]
51	2023	Is Self-Repair a Silver Bullet for Code Generation?	ICLR	[102]
52	2023	LLM4SecHW: Leveraging domain-specific large language model for hardware debugging	AsianHOST	[30]
53	2023	Neural Program Repair with Program Dependence Analysis and Effective Filter Mechanism	arxiv	[179]
54	2023	PATCH: Empowering Large Language Model with Programmer-Intent Guidance and Collaborative-Behavior Simulation for Automatic Bug Fixing	TOSEM	[178]
55	2023	Pre-trained Model-based Automated Software Vulnerability Repair: How Far are We?	TDSC	[174]
56	2023	RAP-Gen: Retrieval-Augmented Patch Generation with CodeT5 for Automatic Program Repair	FSE/ESEC	[138]
57	2023	Refining ChatGPT-Generated Code: Characterizing and Mitigating Code Quality Issues	TOSEM	[86]
58	2023	Repair Is Nearly Generation: Multilingual Program Repair with LLMs	AAAI	[60]
59	2023	Resolving Crash Bugs via Large Language Models: An Empirical Study	arxiv	[20]
60	2023	Retrieval-based prompt selection for code-related few-shot learning	ICSE	[99]
61	2023	Self-Edit: Fault-Aware Code Editor for Code Generation	ACL	[171]
62	2023	SelfEvolve: A Code Evolution Framework via Large Language Models	arxiv	[57]
63	2023	Self-Refine: Iterative Refinement with Self-Feedback	NeurIPS	[92]
64	2023	SkipAnalyzer: An Embodied Agent for Code Analysis with Large Language Models	arxiv	[95]
65	2023	The Best of Both Worlds: Combining Learned Embeddings with Engineered Features for Accurate Prediction of Correct Patches	TOSEM	[128]
66	2023	The Plastic Surgery Hypothesis in the Era of Large Language Models	ASE	[148]
67	2023	The Right Prompts for the Job: Repair Code-Review Defects with Large Language Model	arxiv	[184]
68	2023	Towards Generating Functionally Correct Code Edits from Natural Language Issue Descriptions	ICSE	[23]

Continued on next page

ID	Year	Title	Venue	Reference
69	2023	Training Language Models for Programming Feedback Using Automated Repair Tools	AIED	[67]
70	2023	Vision Transformer-Inspired Automated Vulnerability Repair	TOSEM	[28]
71	2023	What makes good in-context demonstrations for code intelligence tasks with llms?	ASE	[32]
72	2023	ZeroLeak: Using LLMs for Scalable and Cost Effective Side-Channel Patching	ESORICS	[132]
73	2024	A Case Study of LLM for Automated Vulnerability Repair: Assessing Impact of Reasoning and Patch Validation Feedback	AIware	[69]
74	2024	A Deep Dive into Large Language Models for Automated Bug Localization and Repair	FSE/ESEC	[45]
75	2024	A Study of Vulnerability Repair in JavaScript Programs with Large Language Models	WWW	[71]
76	2024	An Empirical Evaluation of Pre-trained Large Language Models for Repairing Declarative Formal Specifications	EMSE	[3]
77	2024	APPT Boosting Automated Patch Correctness Prediction via Pre-trained Language Model	TSE	[173]
78	2024	Assessing the Latent Automated Program Repair Capabilities of Large Language Models using Round-Trip Translation	TOSEM	[117]
79	2024	AutoCodeRover: Autonomous Program Improvement	ISSTA	[180]
80	2024	Automated Program Repair for Introductory Programming Assignments	TLT	[136]
81	2024	Automated Program Repair via Conversation: Fixing 162 out of 337 bugs for \$0.42 each using chatgpt	ISSTA	[151]
82	2024	Automated Test Case Repair Using Language Models	TSE	[119]
83	2024	Benchmarking Automated Program Repair: An Extensive Study on Both Real-World and Artificial Bugs	ISSTA	[107]
84	2024	Chain-of-Thought Prompting of Large Language Models for Discovering and Fixing Software Vulnerabilities	arxiv	[100]
85	2024	CigaR: Cost-efficient Program Repair with LLMs	arxiv	[43]
86	2024	Code repair with llms gives an exploration-exploitation tradeoff	NeurIPS	[125]
87	2024	CodeEditorBench: Evaluating Code Editing Capability of Large Language Models	ICLR	[37]
88	2024	CodeR: Issue Resolving with Multi-Agent and Task Graphs	arxiv	[12]
89	2024	CodeScope: An Execution-based Multilingual Multitask Multidimensional Benchmark for Evaluating LLMs on Code Understanding and Generation	ACL	[156]
90	2024	ConDefects: A New Dataset to Address the Data Leakage Concern for LLM-based Fault Localization and Program Repair	FSE/ESEC	[145]
91	2024	CORE: Resolving Code Quality Issues using LLMs	FSE/ESEC	[135]
92	2024	CraftRTL: High-quality Synthetic Data Generation for Verilog Code Models with Correct-by-Construction Non-Textual Representations and Targeted Code Repair	ICLR	[83]
93	2024	CREF: An LLM-based Conversational Software Repair Framework for Programming Tutors	ISSTA	[158]
94	2024	CYCLE: Learning to Self-Refine the Code Generation	OOPSLA	[18]
95	2024	DebugBench: Evaluating Debugging Capability of Large Language Models	ACL	[131]
96	2024	Divide-and-Conquer: Automating Code Revisions via Localization-and-Revision	TOSEM	[137]
97	2024	Domain Knowledge Matters: Improving Prompts with Fix Templates for Repairing Python Type Errors	ICSE	[112]
98	2024	DrPlanner: Diagnosis and Repair of Motion Planners Using Large Language Models	RAL	[78]
99	2024	Enhancing Automated Program Repair with Solution Design	ASE	[181]

Continued on next page

ID	Year	Title	Venue	Reference
100	2024	Exploring and Lifting the Robustness of LLM-powered Automated Program Repair with Metamorphic Testing	arxiv	[155]
101	2024	Exploring ChatGPT's Capabilities on Vulnerability Management	USENIX Security	[84]
102	2024	Exploring Experiences with Automated Program Repair in Practice	ICSE	[94]
103	2024	Exploring Parameter-Efficient Fine-Tuning of Large Language Model on Automated Program Repair	ASE	[75]
104	2024	Exploring the Potential of ChatGPT in Automated Code Refinement: An Empirical Study	ICSE	[39]
105	2024	FastFixer: An Efficient and Effective Approach for Repairing Programming Assignments	ASE	[81]
106	2024	FlakyFix: Using Large Language Models for Predicting Flaky Test Fix Categories and Test Code Repair	TSE	[25]
107	2024	From Code to Correctness: Closing the Last Mile of Code Generation with Hierarchical Debugging	arxiv	[120]
108	2024	Frustrated with Code Quality Issues? LLMs can Help!	FSE/ESEC	[134]
109	2024	GitBug-Java: A Reproducible Benchmark of Recent Java Bugs	MSR	[122]
110	2024	HDLdebugger: Streamlining HDL debugging with Large Language Models	TODAES	[162]
111	2024	How Far Can We Go with Practical Function-Level Program Repair?	arxiv	[152]
112	2024	Improved Program Repair Methods using Refactoring with GPT Models	SIGCSE TS	[54]
113	2024	Investigating Large Language Models Capabilities for Automatic Code Repair in Python	Cluster Computing	[103]
114	2024	LDB: A Large Language Model Debugger via Verifying Runtime Execution Step by Step	ACL	[186]
115	2024	Leveraging Large Language Model for Automatic Patch Correctness Assessment	TSE	[188]
116	2024	Leveraging Print Debugging to Improve Code Generation in Large Language Models	arxiv	[47]
117	2024	Lost in Translation: A Study of Bugs Introduced by Large Language Models while Translating Code	ICSE	[108]
118	2024	MarsCode Agent: AI-native Automated Bug Fixing	arxiv	[85]
119	2024	MASAI: Modular Architecture for Software-engineering AI Agents	NeurIPS	[6]
120	2024	On Hardware Security Bug Code Fixes By Prompting Large Language Models	TIFS	[1]
121	2024	One Size Does Not Fit All: Multi-granularity Patch Generation for Better Automated Program Repair	ISSTA	[77]
122	2024	OpenCodeInterpreter: Integrating Code Generation with Execution and Refinement	ACL	[185]
123	2024	Out of Context: How important is Local Context in Neural Program Repair?	ICSE	[114]
124	2024	Out of Sight, Out of Mind: Better Automatic Vulnerability Repair by Broadening Input Ranges and Sources	ICSE	[187]
125	2024	Peer-aided Repairer: Empowering Large Language Models to Repair Advanced Student Assignments	arxiv	[182]
126	2024	PyDex: Repairing Bugs in Introductory Python Assignments using LLMs	OOPSLA	[168]
127	2024	PyTy: Repairing Static Type Errors in Python	ICSE	[15]
128	2024	RePair: Automated Program Repair with Process-based Feedback	ACL	[183]
129	2024	RetypeR: Integrated Retrieval-based Automatic Program Repair for Python Type Errors	ICSME	[40]
130	2024	Revisiting Unnaturalness for Automated Program Repair in the Era of Large Language Models	arxiv	[157]
131	2024	RTLFixer: Automatically Fixing RTL Syntax Errors with Large Language Models	DAC	[133]

Continued on next page

ID	Year	Title	Venue	Reference
132	2024	SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering	NeurIPS	[160]
133	2024	SWE-bench: Can Language Models Resolve Real-World GitHub Issues?	ICLR	[58]
134	2024	T5APR: Empowering Automated Program Repair across Languages through Checkpoint Ensemble	JSS	[36]
135	2024	Teaching Large Language Models to Self-Debug	ICLR	[13]
136	2024	Thinkrepair: Self-directed automated program repair	ISSTA	[164]
137	2024	VulAdvisor: Natural Language Suggestion Generation for Software Vulnerability Repair	ASE	[170]
138	2024	When Large Language Models Confront Repository-Level Automatic Program Repair: How Well They Done?	ICSE	[14]
139	2025	A New Era in Software Security: Towards Self-Healing Software via Large Language Models and Formal Verification	AST	[11]
140	2025	A study on Prompt Design, Advantages and Limitations of ChatGPT for Deep Learning Program Repair	AUSE	[10]
141	2025	ACFIX: Guiding LLMs with Mined Common RBAC Practices for Context-Aware Repair of Access Control Vulnerabilities in Smart Contracts	TSE	[172]
142	2025	Adversarial Reasoning for Repair Based on Inferred Program Intent	ISSTA	[163]
143	2025	Alibaba LingmaAgent: Improving Automated Issue Resolution via Comprehensive Repository Exploration	FSE-Companion	[90]
144	2025	Aligning the Objective of LLM-based Program Repair	ICSE	[153]
145	2025	APPATCH: Automated Adaptive Prompting Large Language Models for Real-World Software Vulnerability Patching	USENIX Security	[101]
146	2025	APRMCTS: Improving LLM-based Automated Program Repair with Iterative Tree Search	ASE	[46]
147	2025	Closing the Gap: A User Study on the Real-world Usefulness of AI-powered Vulnerability Detection & Repair in the IDESecurityArtifact-FunctionalArtifact-AvailableArtifact-Reusable	ICSE	[124]
148	2025	Combining Logic and Large Language Models for Assisted Debugging and Repair of ASP Programs	ICST	[9]
149	2025	ContrastRepair: Enhancing Conversation-Based Automated Program Repair via Contrastive Test Case Pairs	TOSEM	[65]
150	2025	Counterexample Guided Program Repair Using Zero-Shot Learning and MaxSAT-based Fault Localization	AAAI	[104]
151	2025	Demystifying LLM-based Software Engineering Agents	FSE/ESEC	[147]
152	2025	Demystifying Memorization in LLM-based Program Repair via a General Hypothesis Testing Framework	FSE/ESEC	[66]
153	2025	DesignRepair: Dual-Stream Design Guideline-Aware Frontend Repair with Large Language Models	ICSE	[166]
154	2025	Error Delayed Is Not Error Handled: Understanding and Fixing Propagated Error-Handling Bugs	FSE/ESEC	[82]
155	2025	Explainable Automated Debugging via Large Language Model-driven Scientific Debugging	EMSE	[61]
156	2025	FixAgent: Hierarchical Multi-Agent Framework for Unified Software Debugging	arxiv	[73]
157	2025	Guiding ChatGPT to Fix Web UI Tests via Explanation-Consistency Checking	ICST	[154]
158	2025	HapRepair: Learn to Repair OpenHarmony Apps	FSE/ESEC	[79]
159	2025	Hierarchical Knowledge Injection for Improving LLM-based Program Repair	ASE	[21]
160	2025	Hybrid Automated Program Repair by Combining Large Language Models and Program Analy	TOSEM	[74]

*Continued on next page*

ID	Year	Title	Venue	Reference
161	2025	Integrating Various Software Artifacts for Better LLM-based Bug Localization and Program Repair	TOSEM	[26]
162	2025	Investigating the Transferability of Code Repair for Low-Resource Programming Languages	NAACL	[142]
163	2025	Knowledge-Enhanced Program Repair for Data Science Code	ICSE	[106]
164	2025	Less is More: Adaptive Program Repair with Bug Localization and Preference Learning	AAAI	[16]
165	2025	MAGIS: LLM-Based Multi-Agent Framework for GitHub Issue Resolution	NeurIPS	[126]
166	2025	MORepair: Teaching LLMs to Repair Code via Multi-Objective Fine-tuning	TOSEM	[159]
167	2025	NIODebugger: A Novel Approach to Repair Non-Idempotent-Outcome Tests with LLM-Based Agent	ICSE	[62]
168	2025	OmniGIRL: A Multilingual and Multimodal Benchmark for GitHub Issue Resolution	ISSTA	[38]
169	2025	OpenHands: An Open Platform for AI Software Developers as Generalist Agents	ICLR	[139]
170	2025	PATCHAGENT: A Practical Program Repair Agent Mimicking Human Expertise	USENIX Security	[165]
171	2025	RAPGen: An Approach for Fixing Code Inefficiencies in Zero-Shot	ICSE	[34]
172	2025	Repair Ingredients Are All You Need: Improving Large Language Model-Based Program Repair via Repair Ingredients Search	ICSE	[169]
173	2025	RepairAgent: An Autonomous, LLM-Based Agent for Program Repair	ICSE	[8]
174	2025	RepairLLaMA: Efficient Representations and Fine-Tuned Adapters for Program Repair	TSE	[121]
175	2025	RepoGraph: Enhancing AI Software Engineering with Repository-level Code Graph	ICLR	[105]
176	2025	RustAssistant: Using LLMs to Fix Compilation Errors in Rust Code	ICSE	[17]
177	2025	Seeing is Fixing: Cross-Modal Reasoning with Multimodal LLMs for Visual Software Issue Fixing	ASE	[52]
178	2025	SWE-bench Multimodal: Do AI Systems Generalize to Visual Software Domains?	ICLR	[161]
179	2025	SWE-GPT: A Process-Centric Language Model for Automated Software Improvement	ISSTA	[89]
180	2025	SWE-RL: Advancing LLM Reasoning via Reinforcement Learning on Open Software Evolution	NeurIPS	[140]
181	2025	SWE-Search: Enhancing Software Agents with Monte Carlo Tree Search and Iterative Refinement	ICLR	[5]
182	2025	SWT-Bench: Testing and Validating Real-World Bug-Fixes with Code Agents	NeurIPS	[97]
183	2025	Teaching AI the ‘Why’ and ‘How’ of Software Vulnerability Fixes	FSE/ESEC	[31]
184	2025	Template-Guided Program Repair in the Era of Large Language Models	ICSE	[51]
185	2025	The Art of Repair: Optimizing Iterative Program Repair with Instruction-Tuned Models	EASE	[118]
186	2025	The Fact Selection Problem in LLM-Based Program Repair	ICSE	[109]
187	2025	The Impact of Fine-tuning Large Language Models on Automated Program Repair	ICSME	[91]
188	2025	Towards Detecting Prompt Knowledge Gaps for Improved LLM-guided Issue Resolution	MSR	[22]
189	2025	When Fine-Tuning LLMs Meets Data Privacy: An Empirical Study of Federated Learning in LLM-Based Program Repair	TOSEM	[88]

## REFERENCES

- [1] Baleegh Ahmad, Shailja Thakur, Benjamin Tan, Ramesh Karri, and Hammond Pearce. 2024. On hardware security bug code fixes by prompting large language models. *IEEE Transactions on Information Forensics and Security* 19 (2024),



4043–4057.

- [2] Toufique Ahmed, Noah Rose Ledesma, and Premkumar Devanbu. 2022. SynShine: Improved Fixing of Syntax Errors. *IEEE Transactions on Software Engineering* 49, 4 (2022), 2169–2181.
- [3] Mohammadh Alhanahnah, Md Rashedul Hasan, Lisong Xu, and Hamid Bagheri. 2025. An empirical evaluation of pre-trained large language models for repairing declarative formal specifications. *Empirical Software Engineering* 30, 5 (2025), 1–38.
- [4] Kamel Alrashedy and Abdullah Aljasser. 2023. Can LLMs Patch Security Issues? *arXiv preprint arXiv:2312.00024* (2023).
- [5] Antonis Antoniadis, Albert Örwall, Kexun Zhang, Yuxi Xie, Anirudh Goyal, and William Wang. 2024. SWE-Search: Enhancing Software Agents with Monte Carlo Tree Search and Iterative Refinement. *arXiv preprint arXiv:2410.20285* (2024).
- [6] Daman Arora, Atharv Sonwane, Nalin Wadhwa, Abhav Mehrotra, Saiteja Utpala, Ramakrishna Bairi, Aditya Kanade, and Nagarajan Natarajan. 2024. Masai: Modular architecture for software-engineering ai agents. *arXiv preprint arXiv:2406.11638* (2024).
- [7] Berkay Berabi, Jingxuan He, Veselin Raychev, and Martin Vechev. 2021. TFix: Learning to Fix Coding Errors with a Text-to-Text Transformer. In *International Conference on Machine Learning*. PMLR, 780–791.
- [8] Islem Bouzenia, Premkumar Devanbu, and Michael Pradel. 2024. RepairAgent: An Autonomous, LLM-Based Agent for Program Repair. *arXiv preprint arXiv:2403.17134* (2024).
- [9] Ricardo Brancas, Vasco Manquinho, and Ruben Martins. 2025. Combining Logic and Large Language Models for Assisted Debugging and Repair of ASP Programs. In *2025 IEEE Conference on Software Testing, Verification and Validation (ICST)*. IEEE, 646–657.
- [10] Jialun Cao, Meiziniu Li, Ming Wen, and Shing-chi Cheung. 2023. A Study on Prompt Design, Advantages and Limitations of ChatGPT for Deep Learning Program Repair. *arXiv preprint arXiv:2304.08191* (2023).
- [11] Yiannis Charalambous, Norbert Tihanyi, Ridhi Jain, Youcheng Sun, Mohamed Amine Ferrag, and Lucas C Cordeiro. 2023. A New Era in Software Security: Towards Self-Healing Software Via Large Language Models and Formal Verification. *arXiv preprint arXiv:2305.14752* (2023).
- [12] Dong Chen, Shaoxin Lin, Muhan Zeng, Daoguang Zan, Jian-Gang Wang, Anton Cheshkov, Jun Sun, Hao Yu, Guoliang Dong, Artem Aliev, et al. 2024. Coder: Issue resolving with multi-agent and task graphs. *arXiv preprint arXiv:2406.01304* (2024).
- [13] Xinyun Chen, Maxwell Lin, Nathanael Schärli, and Denny Zhou. 2024. Teaching Large Language Models to Self-Debug. In *The Twelfth International Conference on Learning Representations*. <https://openreview.net/forum?id=KuPixIqPiq>
- [14] Yuxiao Chen, Jingzheng Wu, Xiang Ling, Changjiang Li, Zhiqing Rui, Tianyue Luo, and Yanjun Wu. 2024. When Large Language Models Confront Repository-Level Automatic Program Repair: How Well They Done?. In *Proceedings of the 2024 IEEE/ACM 46th International Conference on Software Engineering: Companion Proceedings*. 459–471.
- [15] Yiu Wai Chow, Luca Di Grazia, and Michael Pradel. 2024. PyTy: Repairing Static Type Errors in Python. In *Proceedings of the 46th International Conference on Software Engineering*. 871–871.
- [16] Zhenlong Dai, Bingrui Chen, Zhuoluo Zhao, Xiu Tang, Sai Wu, Chang Yao, Zhipeng Gao, and Jingyuan Chen. 2025. Less is More: Adaptive Program Repair with Bug Localization and Preference Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 128–136.
- [17] Pantazis Deligiannis, Akash Lal, Nikita Mehrotra, and Aseem Rastogi. 2023. Fixing Rust Compilation Errors Using LLMs. *arXiv preprint arXiv:2308.05177* (2023).
- [18] Yangruibo Ding, Marcus J Min, Gail Kaiser, and Baishakhi Ray. 2024. CYCLE: Learning to Self-Refine the Code Generation. *arXiv preprint arXiv:2403.18746* (2024).
- [19] Dawn Drain, Chen Wu, Alexey Svyatkovskiy, and Neel Sundaresan. 2021. Generating Bug-Fixes Using Pretrained Transformers. In *Proceedings of the 5th ACM SIGPLAN International Symposium on Machine Programming*. 1–8.
- [20] Xueying Du, Mingwei Liu, Juntao Li, Hanlin Wang, Xin Peng, and Yiling Lou. 2023. Resolving Crash Bugs Via Large Language Models: An Empirical Study. *arXiv preprint arXiv:2312.10448* (2023).
- [21] Ramtin Ehsani, Esteban Parra, Sonia Haiduc, and Preetha Chatterjee. 2025. Bug Fixing with Broader Context: Enhancing LLM-Based Program Repair via Layered Knowledge Injection. *arXiv preprint arXiv:2506.24015* (2025).
- [22] Ramtin Ehsani, Sakshi Pathak, and Preetha Chatterjee. 2025. Towards detecting prompt knowledge gaps for improved llm-guided issue resolution. In *2025 IEEE/ACM 22nd International Conference on Mining Software Repositories (MSR)*. IEEE, 699–711.
- [23] Sarah Fakhoury, Saikat Chakraborty, Madan Musuvathi, and Shuvendu K Lahiri. 2023. Towards Generating Functionally Correct Code Edits from Natural Language Issue Descriptions. *arXiv preprint arXiv:2304.03816* (2023).
- [24] Zhiyu Fan, Xiang Gao, Martin Mirchev, Abhik Roychoudhury, and Shin Hwei Tan. 2023. Automated Repair of Programs from Large Language Models. In *2023 IEEE/ACM 45th International Conference on Software Engineering*. IEEE, 1469–1481.

- [25] Sakina Fatima, Hadi Hemmati, and Lionel C Briand. 2024. FlakyFix: Using large language models for predicting flaky test fix categories and test code repair. *IEEE Transactions on Software Engineering* 50, 12 (2024), 3146–3171.
- [26] Qiong Feng, Xiaotian Ma, Jiayi Sheng, Ziyuan Feng, Wei Song, and Peng Liang. 2025. Integrating Various Software Artifacts for Better LLM-based Bug Localization and Program Repair. *ACM Transactions on Software Engineering and Methodology* (2025).
- [27] Emily First, Markus Rabe, Talia Ringer, and Yuriy Brun. 2023. Baldur: Whole-proof generation and repair with large language models. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 1229–1241.
- [28] Michael Fu, Van Nguyen, Chakkrit Tantithamthavorn, Dinh Phung, and Trung Le. 2023. Vision Transformer-Inspired Automated Vulnerability Repair. *ACM Transactions on Software Engineering and Methodology* (2023).
- [29] Michael Fu, Chakkrit Tantithamthavorn, Trung Le, Van Nguyen, and Phung Dinh. 2022. VulRepair: A T5-Based Automated Software Vulnerability Repair. In *the ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. ACM, 935–947.
- [30] Weimin Fu, Kaichen Yang, Raj Gautam Dutta, Xiaolong Guo, and Gang Qu. 2023. LLM4SecHW: Leveraging domain-specific large language model for hardware debugging. In *2023 Asian Hardware Oriented Security and Trust Symposium (AsianHOST)*. IEEE, 1–6.
- [31] Amiao Gao, Zenong Zhang, Simin Wang, Liguang Huang, Shiyi Wei, and Vincent Ng. 2025. Teaching AI the ‘Why’ and ‘How’ of Software Vulnerability Fixes. *Proceedings of the ACM on Software Engineering* 2, FSE (2025), 2006–2029.
- [32] Shuzheng Gao, Xin-Cheng Wen, Cuiyun Gao, Wenxuan Wang, Hongyu Zhang, and Michael R Lyu. 2023. What Makes Good In-Context Demonstrations for Code Intelligence Tasks with LLMs?. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering*. IEEE, 761–773.
- [33] Spandan Garg, Roshanak Zilouchian Moghaddam, Colin B Clement, Neel Sundaresan, and Chen Wu. 2022. DeepDevPERF: A Deep Learning-Based Approach for Improving Software Performance. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 948–958.
- [34] Spandan Garg, Roshanak Zilouchian Moghaddam, and Neel Sundaresan. 2023. Rapgen: An Approach for Fixing Code Inefficiencies in Zero-Shot. *arXiv preprint arXiv:2306.17077* (2023).
- [35] Haotong Ge and Yuemeng Wu. 2023. An Empirical Study of Adoption of ChatGPT for Bug Fixing among Professional Developers. *Innovation & Technology Advances* 1, 1 (Jun. 2023), 21–29. <https://doi.org/10.61187/ita.v1i1.19>
- [36] Reza Gharibi, Mohammad Hadi Sadreddini, and Seyed Mostafa Fakhrahmad. 2024. T5APR: Empowering automated program repair across languages through checkpoint ensemble. *Journal of Systems and Software* 214 (2024), 112083.
- [37] Jiawei Guo, Ziming Li, Xueling Liu, Kaijing Ma, Tianyu Zheng, Zhouliang Yu, Ding Pan, Yizhi Li, Ruibo Liu, Yue Wang, et al. 2024. CodeEditorBench: Evaluating Code Editing Capability of Large Language Models. *arXiv preprint arXiv:2404.03543* (2024).
- [38] Lianghong Guo, Wei Tao, Runhan Jiang, Yanlin Wang, Jiachi Chen, Xilin Liu, Yuchi Ma, Mingzhi Mao, Hongyu Zhang, and Zibin Zheng. 2025. OmniGIRL: A Multilingual and Multimodal Benchmark for GitHub Issue Resolution. *Proceedings of the ACM on Software Engineering* 2, ISSTA (2025), 24–46.
- [39] Qi Guo, Junming Cao, Xiaofei Xie, Shangqing Liu, Xiaohong Li, Bihuan Chen, and Xin Peng. 2023. Exploring the Potential of ChatGPT in Automated Code Refinement: An Empirical Study. In *2024 IEEE/ACM 46th International Conference on Software Engineering (ICSE)*. IEEE Computer Society, 379–391.
- [40] Sichong Hao, Xianjun Shi, and Hongwei Liu. 2024. RetypeR: Integrated Retrieval-based Automatic Program Repair for Python Type Errors. In *2024 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. IEEE, 199–210.
- [41] Sichong Hao, Xianjun Shi, Hongwei Liu, and Yanjun Shu. 2023. Enhancing Code Language Models for Program Repair by Curricular Fine-tuning Framework. In *2023 IEEE International Conference on Software Maintenance and Evolution (ICSME)*. IEEE, 136–146.
- [42] Md Mahim Anjum Haque, Wasi Uddin Ahmad, Ismini Lourentzou, and Chris Brown. 2023. FixEval: Execution-Based Evaluation of Program Fixes for Programming Problems. In *2023 IEEE/ACM International Workshop on Automated Program Repair*. IEEE, 11–18.
- [43] Dávid Hidvégi, Khashayar Ettemadi, Sofia Bobadilla, and Martin Monperrus. 2024. CigaR: Cost-efficient Program Repair with LLMs. *arXiv preprint arXiv:2402.06598* (2024).
- [44] Dániel Horváth, Viktor Csuik, Tibor Gyimóthy, and László Vidács. 2023. An Extensive Study on Model Architecture and Program Representation in the Domain of Learning-Based Automated Program Repair. In *2023 IEEE/ACM International Workshop on Automated Program Repair (APR)*. 31–38.
- [45] Soneya Binta Hossain, Nan Jiang, Qiang Zhou, Xiaopeng Li, Wen-Hao Chiang, Yingjun Lyu, Hoan Nguyen, and Omer Tripp. 2024. A Deep Dive into Large Language Models for Automated Bug Localization and Repair. *Proceedings of the ACM on Software Engineering* 1, FSE (2024), 1471–1493.



- [46] Haichuan Hu, Congqing He, Hao Zhang, Xiaochen Xie, and Qunjun Zhang. 2025. APRMCTS: Improving LLM-based Automated Program Repair with Iterative Tree Search. *arXiv preprint arXiv:2507.01827* (2025).
- [47] Xueyu Hu, Kun Kuang, Jiankai Sun, Hongxia Yang, and Fei Wu. 2024. Leveraging Print Debugging to Improve Code Generation in Large Language Models. *arXiv preprint arXiv:2401.05319* (2024).
- [48] Yaojie Hu, Xingjian Shi, Qiang Zhou, and Lee Pike. 2022. Fix Bugs with Transformer through a Neural-Symbolic Edit Grammar. In *Deep Learning for Code Workshop*. [https://openreview.net/forum?id=SBgE6i\\_WkZq](https://openreview.net/forum?id=SBgE6i_WkZq)
- [49] Dong Huang, Qingwen Bu, Jie M Zhang, Michael Luck, and Heming Cui. 2023. AgentCoder: Multi-Agent-based Code Generation with Iterative Testing and Optimisation. *arXiv preprint arXiv:2312.13010* (2023).
- [50] Kai Huang, Jian Zhang, Xinlei Bao, Xu Wang, and Yang Liu. 2025. Comprehensive Fine-Tuning Large Language Models of Code for Automated Program Repair. *IEEE Transactions on Software Engineering* 51, 4 (2025), 904–928.
- [51] Kai Huang, Jian Zhang, Xiangxin Meng, and Yang Liu. 2025. Template-Guided Program Repair in the Era of Large Language Models. In *ICSE*. 1895–1907.
- [52] Kai Huang, Jian Zhang, Xiaofei Xie, and Chunyang Chen. 2025. Seeing is Fixing: Cross-Modal Reasoning with Multimodal LLMs for Visual Software Issue Fixing. *arXiv preprint arXiv:2506.16136* (2025).
- [53] Qing Huang, Jiahui Zhu, Zhenchang Xing, Huan Jin, Changjing Wang, and Xiwei Xu. 2023. A Chain of AI-Based Solutions for Resolving Fqns and Fixing Syntax Errors in Partial Code. *arXiv preprint arXiv:2306.11981* (2023).
- [54] Ryosuke Ishizue, Kazunori Sakamoto, Hironori Washizaki, and Yoshiaki Fukazawa. 2024. Improved Program Repair Methods using Refactoring with GPT Models. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 1*. 569–575.
- [55] Nan Jiang, Kevin Liu, Thibaud Lutellier, and Lin Tan. 2023. Impact of Code Language Models on Automated Program Repair. In *Proceedings of the 45th International Conference on Software Engineering*. 1430–1442.
- [56] Nan Jiang, Thibaud Lutellier, and Lin Tan. 2021. CURE: Code-Aware Neural Machine Translation for Automatic Program Repair. In *Proceedings of the 43rd IEEE/ACM International Conference on Software Engineering*. 1161–1173.
- [57] Shuyang Jiang, Yuhao Wang, and Yu Wang. 2023. SelfEvolve: A Code Evolution Framework via Large Language Models. *arXiv preprint arXiv:2306.02907* (2023).
- [58] Carlos E Jimenez, John Yang, Alexander Wettig, Shunyu Yao, Kexin Pei, Ofir Press, and Karthik R Narasimhan. 2024. SWE-bench: Can Language Models Resolve Real-world Github Issues?. In *The Twelfth International Conference on Learning Representations*. <https://openreview.net/forum?id=VTF8yNQM66>
- [59] Matthew Jin, Syed Shahriar, Michele Tufano, Xin Shi, Shuai Lu, Neel Sundaresan, and Alexey Svyatkovskiy. 2023. InferFix: End-to-End Program Repair with LLMs. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. ACM, 1646–1656.
- [60] Harshit Joshi, José Cambronero Sanchez, Sumit Gulwani, Vu Le, Gust Verbruggen, and Ivan Radiček. 2023. Repair Is Nearly Generation: Multilingual Program Repair with LLMs. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 37. 5131–5140.
- [61] Sungmin Kang, Bei Chen, Shin Yoo, and Jian-Guang Lou. 2023. Explainable Automated Debugging Via Large Language Model-Driven Scientific Debugging. *arXiv preprint arXiv:2304.02195* (2023).
- [62] Kaiyao Ke. 2025. NIODebugger: A Novel Approach to Repair Non-Idempotent-Outcome Tests with LLM-Based Agent. In *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. IEEE Computer Society, 762–762.
- [63] Misoo Kim, Youngkyoung Kim, Hohyeon Jeong, Jinseok Heo, Sungoh Kim, Hyunhee Chung, and Eunseok Lee. 2022. An Empirical Study of Deep Transfer Learning-Based Program Repair for Kotlin Projects. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 1441–1452.
- [64] Sophia D Kolak, Ruben Martins, Claire Le Goues, and Vincent Josua Hellendoorn. 2022. Patch Generation with Language Models: Feasibility and Scaling Behavior. In *International Conference on Learning Representations Deep Learning for Code Workshop*.
- [65] Jiaolong Kong, Mingfei Cheng, Xiaofei Xie, Shangqing Liu, Xiaoning Du, and Qi Guo. 2024. ContrastRepair: Enhancing Conversation-Based Automated Program Repair Via Contrastive Test Case Pairs. *arXiv preprint arXiv:2403.01971* (2024).
- [66] Jiaolong Kong, Xiaofei Xie, and Shangqing Liu. 2025. Demystifying Memorization in LLM-Based Program Repair via a General Hypothesis Testing Framework. *Proceedings of the ACM on Software Engineering* 2, FSE (2025), 2712–2734.
- [67] Charles Koutchme. 2023. Training Language Models for Programming Feedback Using Automated Repair Tools. In *International Conference on Artificial Intelligence in Education*. Springer, 830–835.
- [68] Charles Koutchme, Sami Sarsa, Juho Leinonen, Arto Hellas, and Paul Denny. 2023. Automated Program Repair Using Generative Models for Code Infilling. In *International Conference on Artificial Intelligence in Education*. Springer, 798–803.
- [69] Ummay Kulsum, Haotian Zhu, Bowen Xu, and Marcelo d’Amorim. 2024. A Case Study of LLM for Automated Vulnerability Repair: Assessing Impact of Reasoning and Patch Validation Feedback. In *Proceedings of the 1st ACM*

*International Conference on AI-Powered Software*. 103–111.

- [70] Márk Lajkó, Viktor Csuvi, and László Vidács. 2022. Towards Javascript Program Repair with Generative Pre-trained Transformer. In *2022 IEEE/ACM International Workshop on Automated Program Repair*. IEEE, 61–68.
- [71] Tan Khang Le, Saba Alimadadi, and Steven Y Ko. 2024. A Study of Vulnerability Repair in JavaScript Programs with Large Language Models. *arXiv preprint arXiv:2403.13193* (2024).
- [72] Thanh Le-Cong, Duc-Minh Luong, Xuan Bach D Le, David Lo, Nhat-Hoa Tran, Bui Quang-Huy, and Quyet-Thang Huynh. 2023. Invalidator: Automated Patch Correctness Assessment Via Semantic and Syntactic Reasoning. *IEEE Transactions on Software Engineering* 49, 06 (2023), 3411–3429.
- [73] Cheryl Lee, Chunqiu Steven Xia, Longji Yang, Jen-tse Huang, Zhouruixin Zhu, Lingming Zhang, and Michael R Lyu. 2024. A unified debugging approach via llm-based multi-agent synergy. *arXiv preprint arXiv:2404.17153* (2024).
- [74] Fengjie Li, Jiajun Jiang, Jiajun Sun, and Hongyu Zhang. 2025. Hybrid automated program repair by combining large language models and program analysis. *ACM Transactions on Software Engineering and Methodology* 34, 7 (2025), 1–28.
- [75] Guochang Li, Chen Zhi, Jialiang Chen, Junxiao Han, and Shuiguang Deng. 2024. Exploring Parameter-Efficient Fine-Tuning of Large Language Model on Automated Program Repair. In *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 719–731.
- [76] Yi Li, Shaohua Wang, and Tien N. Nguyen. 2022. DEAR: A Novel Deep Learning-based Approach for Automated Program Repair. In *Proceedings of the 44th International Conference on Software Engineering*. 511–523.
- [77] Bo Lin, Shangwen Wang, Ming Wen, Liqian Chen, and Xiaoguang Mao. 2024. One Size Does Not Fit All: Multi-granularity Patch Generation for Better Automated Program Repair. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 1554–1566.
- [78] Yuanfei Lin, Chenran Li, Mingyu Ding, Masayoshi Tomizuka, Wei Zhan, and Matthias Althoff. 2024. DrPlanner: Diagnosis and Repair of Motion Planners Using Large Language Models. *arXiv preprint arXiv:2403.07470* (2024).
- [79] Zhihao Lin, Mingyi Zhou, Wei Ma, Chi Chen, Yun Yang, Jun Wang, Chunming Hu, and Li Li. 2025. HapRepair: Learn to Repair OpenHarmony Apps. In *Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering*. 319–330.
- [80] Changshu Liu, Pelin Cetin, Yogesh Patodia, Saikat Chakraborty, Yangruibo Ding, and Baishakhi Ray. 2023. Automated Code Editing with Search-Generate-Modify. *arXiv preprint arXiv:2306.06490* (2023).
- [81] Fang Liu, Zhenwei Liu, Qianhui Zhao, Jing Jiang, Li Zhang, Zian Sun, Ge Li, Zhongqi Li, and Yuchi Ma. 2024. FastFixer: an efficient and effective approach for repairing programming assignments. In *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 669–680.
- [82] Haoran Liu, Shanshan Li, Zhouyang Jia, Yuanliang Zhang, Linxiao Bai, Si Zheng, Xiaoguang Mao, and Xiangke Liao. 2025. Error Delayed Is Not Error Handled: Understanding and Fixing Propagated Error-Handling Bugs. *Proceedings of the ACM on Software Engineering* 2, FSE (2025), 2571–2593.
- [83] Mingjie Liu, Yun-Da Tsai, Wenfei Zhou, and Haoxing Ren. 2024. CraftRTL: High-quality Synthetic Data Generation for Verilog Code Models with Correct-by-Construction Non-Textual Representations and Targeted Code Repair. *arXiv preprint arXiv:2409.12993* (2024).
- [84] Peiyu Liu, Junming Liu, Lirong Fu, Kangjie Lu, Yifan Xia, Xuhong Zhang, Wenzhi Chen, Haiqin Weng, Shouling Ji, and Wenhai Wang. 2023. How ChatGPT Is Solving Vulnerability Management Problem. *arXiv preprint arXiv:2311.06530* (2023).
- [85] Yizhou Liu, Pengfei Gao, Xincheng Wang, Jie Liu, Yexuan Shi, Zhao Zhang, and Chao Peng. 2024. MarsCode Agent: AI-native Automated Bug Fixing. *arXiv preprint arXiv:2409.00899* (2024).
- [86] Yue Liu, Thanh Le-Cong, Ratnadira Widayarsi, Chakkrit Tantithamthavorn, Li Li, Xuan-Bach D. Le, and David Lo. 2024. Refining ChatGPT-Generated Code: Characterizing and Mitigating Code Quality Issues. *ACM Transactions on Software Engineering and Methodology* (2024). <https://doi.org/10.1145/3643674> Just Accepted.
- [87] Vadim Liventsev, Anastasiia Grishina, Aki Härmä, and Leon Moonen. 2023. Fully Autonomous Programming with Large Language Models. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '23)*. 1146–1155.
- [88] Wenqiang Luo, Jacky Keung, Boyang Yang, He Ye, Claire Le Goues, Tegawende F Bissyande, Haoye Tian, and Xuan Bach D Le. 2024. When Fine-Tuning LLMs Meets Data Privacy: An Empirical Study of Federated Learning in LLM-Based Program Repair. *ACM Transactions on Software Engineering and Methodology* (2024).
- [89] Yingwei Ma, Rongyu Cao, Yongchang Cao, Yue Zhang, Jue Chen, Yibo Liu, Yuchen Liu, Binhua Li, Fei Huang, and Yongbin Li. 2025. SWE-GPT: A Process-Centric Language Model for Automated Software Improvement. *Proceedings of the ACM on Software Engineering* 2, ISSTA (2025), 2362–2383.
- [90] Yingwei Ma, Qingping Yang, Rongyu Cao, Binhua Li, Fei Huang, and Yongbin Li. 2025. Alibaba lingmaagent: Improving automated issue resolution via comprehensive repository exploration. In *Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering*. 238–249.

- [91] Roman Macháček, Anastasiia Grishina, Max Hort, and Leon Moonen. 2025. The Impact of Fine-tuning Large Language Models on Automated Program Repair. *arXiv preprint arXiv:2507.19909* (2025).
- [92] Aman Madaan, Niket Tandon, Prakhar Gupta, Skyler Hallinan, Luyu Gao, Sarah Wiegrefe, Uri Alon, Nouha Dziri, Shrimai Prabhumoye, Yiming Yang, Shashank Gupta, Bodhisattwa Prasad Majumder, Katherine Hermann, Sean Welleck, Amir Yazdanbakhsh, and Peter Clark. 2023. Self-Refine: Iterative Refinement with Self-Feedback. In *Thirty-seventh Conference on Neural Information Processing Systems*. <https://openreview.net/forum?id=S37hOerQLB>
- [93] Ehsan Mashhadi and Hadi Hemmati. 2021. Applying Codebert for Automated Program Repair of Java Simple Bugs. In *Proceedings Companion of the 18th IEEE/ACM International Conference on Mining Software Repositories (MSR'21)*. 505–509.
- [94] Fairuz Nawer Meem, Justin Smith, and Brittany Johnson. 2024. Exploring Experiences with Automated Program Repair in Practice. In *2024 IEEE/ACM 46th International Conference on Software Engineering (ICSE)*. IEEE Computer Society, 870–870.
- [95] Mohammad Mahdi Mohajer, Reem Aleithan, Nima Shiri Harzevili, Moshi Wei, Alvine Boaye Belle, Hung Viet Pham, and Song Wang. 2023. SkipanAlyzer: An Embodied Agent for Code Analysis with Large Language Models. *arXiv preprint arXiv:2310.18532* (2023).
- [96] Seungjun Moon, Yongho Song, Hyungjoo Chae, Dongjin Kang, Taeyoon Kwon, Kai Tzu-iunn Ong, Seung-won Hwang, and Jinyoung Yeo. 2023. Coffee: Boost Your Code Llm by Fixing Bugs with Feedback. *arXiv preprint arXiv:2311.07215* (2023).
- [97] Niels Mündler, Mark Müller, Jingxuan He, and Martin Vechev. 2024. SWT-bench: Testing and validating real-world bug-fixes with code agents. *Advances in Neural Information Processing Systems* 37 (2024), 81857–81887.
- [98] Emanuele Antonio Napoli and Valentina Gatteschi. 2023. Evaluating ChatGPT for Smart Contracts Vulnerability Correction. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 1828–1833.
- [99] Noor Nashid, Mifta Sintaha, and Ali Mesbah. 2023. Retrieval-Based Prompt Selection for Code-Related Few-Shot Learning. In *Proceedings of the 45th International Conference on Software Engineering*. IEEE, 2450–2462.
- [100] Yu Nong, Mohammed Aldeen, Long Cheng, Hongxin Hu, Feng Chen, and Haipeng Cai. 2024. Chain-of-Thought Prompting of Large Language Models for Discovering and Fixing Software Vulnerabilities. *arXiv preprint arXiv:2402.17230* (2024).
- [101] Yu Nong, Haoran Yang, Long Cheng, Hongxin Hu, and Haipeng Cai. 2025. {APPATCH}: Automated Adaptive Prompting Large Language Models for {Real-World} Software Vulnerability Patching. In *34th USENIX Security Symposium (USENIX Security 25)*. 4481–4500.
- [102] Theo X. Olausson, Jeevana Priya Inala, Chenglong Wang, Jianfeng Gao, and Armando Solar-Lezama. 2024. Is Self-Repair a Silver Bullet for Code Generation?. In *The Twelfth International Conference on Learning Representations*. <https://openreview.net/forum?id=y0GJXRungR>
- [103] Safwan Omari, Kshitiz Basnet, and Mohammad Wardat. 2024. Investigating large language models capabilities for automatic code repair in python. *Cluster Computing* 27, 8 (2024), 10717–10731.
- [104] Pedro Orvalho, Mikoláš Janota, and Vasco M Manquinho. 2025. Counterexample Guided Program Repair Using Zero-Shot Learning and MaxSAT-based Fault Localization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 649–657.
- [105] Siru Ouyang, Wenhao Yu, Kaixin Ma, Zilin Xiao, Zhihan Zhang, Mengzhao Jia, Jiawei Han, Hongming Zhang, and Dong Yu. 2024. RepoGraph: Enhancing AI Software Engineering with Repository-level Code Graph. *arXiv preprint arXiv:2410.14684* (2024).
- [106] Shuyin Ouyang, Jie M Zhang, Zeyu Sun, and Albert Merono Penuela. 2025. Knowledge-Enhanced Program Repair for Data Science Code. In *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. IEEE, 898–910.
- [107] Yicheng Ouyang, Jun Yang, and Lingming Zhang. 2024. Benchmarking automated program repair: An extensive study on both real-world and artificial bugs. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 440–452.
- [108] Rangeet Pan, Ali Reza Ibrahimzada, Rahul Krishna, Divya Sankar, Lambert Pouguem Wassi, Michele Merler, Boris Sobolev, Raju Pavuluri, Saurabh Sinha, and Reyhaneh Jabbarvand. 2024. Lost in Translation: A Study of Bugs Introduced by Large Language Models While Translating Code. In *2024 IEEE/ACM 46th International Conference on Software Engineering*. IEEE Computer Society, 866–866.
- [109] Nikhil Parasaram, Huijie Yan, Boyu Yang, Zineb Flahy, Abriele Qudsi, Damian Ziaber, Earl T Barr, and Sergey Mechtaev. 2025. The Fact Selection Problem in LLM-Based Program Repair. In *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. IEEE, 2574–2586.
- [110] Rishov Paul, Md Mohib Hossain, Mohammed Latif Siddiq, Masum Hasan, Anindya Iqbal, and Joanna CS Santos. 2023. Enhancing Automated Program Repair through Fine-Tuning and Prompt Engineering. *arXiv preprint arXiv:2304.07840* (2023).

- [111] Hammond Pearce, Benjamin Tan, Baleegh Ahmad, Ramesh Karri, and Brendan Dolan-Gavitt. 2023. Examining Zero-Shot Vulnerability Repair with Large Language Models. In *2023 IEEE Symposium on Security and Privacy*. IEEE, 2339–2356.
- [112] Yun Peng, Shuzheng Gao, Cuiyun Gao, Yintong Huo, and Michael Lyu. 2024. Domain Knowledge Matters: Improving Prompts with Fix Templates for Repairing Python Type Errors. In *Proceedings of the 46th IEEE/ACM International Conference on Software Engineering*. 1–13.
- [113] Julian Aron Prenner, Hlib Babii, and Romain Robbes. 2022. Can OpenAI’s Codex Fix Bugs? An Evaluation on QuixBugs. In *Proceedings of the Third International Workshop on Automated Program Repair*. 69–75.
- [114] Julian Aron Prenner and Romain Robbes. 2023. Out of Context: How Important Is Local Context in Neural Program Repair? *arXiv preprint arXiv:2312.04986* (2023).
- [115] Francisco Ribeiro, Rui Abreu, and João Saraiva. 2022. Framing Program Repair as Code Completion. In *Proceedings of the Third International Workshop on Automated Program Repair*. IEEE, 38–45.
- [116] Francisco Ribeiro, José Nuno Castro de Macedo, Kanae Tsushima, Rui Abreu, and João Saraiva. 2023. GPT-3-Powered Type Error Debugging: Investigating the Use of Large Language Models for Code Repair. In *Proceedings of the 16th ACM SIGPLAN International Conference on Software Language Engineering*. 111–124.
- [117] Fernando Vallecillos Ruiz, Anastasiia Grishina, Max Hort, and Leon Moonen. 2024. A Novel Approach for Automatic Program Repair Using Round-Trip Translation with Large Language Models. *arXiv preprint arXiv:2401.07994* (2024).
- [118] Fernando Vallecillos Ruiz, Max Hort, and Leon Moonen. 2025. The Art of Repair: Optimizing Iterative Program Repair with Instruction-Tuned Models. *arXiv preprint arXiv:2505.02931* (2025).
- [119] Ahmadreza Saboor Yaraghi, Darren Holden, Nafiseh Kahani, and Lionel Briand. 2024. Automated Test Case Repair Using Language Models. *arXiv e-prints* (2024), arXiv–2401.
- [120] Yuling Shi, Songsong Wang, Chengcheng Wan, Min Wang, and Xiaodong Gu. 2024. From code to correctness: Closing the last mile of code generation with hierarchical debugging. *arXiv preprint arXiv:2410.01215* (2024).
- [121] André Silva, Sen Fang, and Martin Monperrus. 2023. RepairLLaMA: Efficient Representations and Fine-Tuned Adapters for Program Repair. *arXiv preprint arXiv:2312.15698* (2023).
- [122] André Silva, Nuno Saavedra, and Martin Monperrus. 2024. GitBug-Java: A Reproducible Benchmark of Recent Java Bugs. *arXiv preprint arXiv:2402.02961* (2024).
- [123] Dominik Sobania, Martin Briesch, Carol Hanna, and Justyna Petke. 2023. An Analysis of the Automatic Bug Fixing Performance of ChatGPT. In *2023 IEEE/ACM International Workshop on Automated Program Repair*. 23–30.
- [124] Benjamin Steenhoeck, Kalpathy Sivaraman, Renata Saldivar Gonzalez, Yevhen Mohylevskyy, Roshanak Zilouchian Moghaddam, and Wei Le. 2025. Closing the Gap: A User Study on the Real-world Usefulness of AI-powered Vulnerability Detection & Repair in the IDE. In *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. IEEE Computer Society, 663–663.
- [125] Hao Tang, Keya Hu, Jin Zhou, Si Cheng Zhong, Wei-Long Zheng, Xujie Si, and Kevin Ellis. 2024. Code Repair with LLMs gives an Exploration-Exploitation Tradeoff. *Advances in Neural Information Processing Systems* 37 (2024), 117954–117996.
- [126] Wei Tao, Yucheng Zhou, Yanlin Wang, Wenqiang Zhang, Hongyu Zhang, and Yu Cheng. 2024. MAGIS: LLM-Based Multi-Agent Framework for GitHub Issue Resolution. *Advances in Neural Information Processing Systems* 37 (2024), 51963–51993.
- [127] Haoye Tian, Kui Liu, Abdoul Kader Kaboré, Anil Koyuncu, Li Li, Jacques Klein, and Tegawendé F Bissyandé. 2020. Evaluating Representation Learning of Code Changes for Predicting Patch Correctness in Program Repair. In *Proceedings of the 35th IEEE/ACM International Conference on Automated Software Engineering*. 981–992.
- [128] Haoye Tian, Kui Liu, Yinghua Li, Abdoul Kader Kaboré, Anil Koyuncu, Andrew Habib, Li Li, Junhao Wen, Jacques Klein, and Tegawendé F Bissyandé. 2022. The Best of Both Worlds: Combining Learned Embeddings with Engineered Features for Accurate Prediction of Correct Patches. *ACM Transactions on Software Engineering and Methodology* 1, 1 (2022), 1–1.
- [129] Haoye Tian, Weiqi Lu, Tsz On Li, Xunzhu Tang, Shing-Chi Cheung, Jacques Klein, and Tegawendé F Bissyandé. 2023. Is ChatGPT the Ultimate Programming Assistant—How Far Is It? *arXiv preprint arXiv:2304.11938* (2023).
- [130] Haoye Tian, Xunzhu Tang, Andrew Habib, Shangwen Wang, Kui Liu, Xin Xia, Jacques Klein, and Tegawendé F Bissyandé. 2022. Is This Change the Answer to That Problem? Correlating Descriptions of Bug and Code Changes for Evaluating Patch Correctness. In *37th IEEE/ACM International Conference on Automated Software Engineering*. IEEE, 1–13.
- [131] Runchu Tian, Yining Ye, Yujia Qin, Xin Cong, Yankai Lin, Zhiyuan Liu, and Maosong Sun. 2024. DebugBench: Evaluating Debugging Capability of Large Language Models. *arXiv preprint arXiv:2401.04621* (2024).
- [132] M Caner Tol and Berk Sunar. 2023. Zeroleak: Using Llms for Scalable and Cost Effective Side-Channel Patching. *arXiv preprint arXiv:2308.13062* (2023).

- [133] YunDa Tsai, Mingjie Liu, and Haoxing Ren. 2023. RTLFixer: Automatically Fixing RTL Syntax Errors with Large Language Models. *arXiv preprint arXiv:2311.16543* (2023).
- [134] Nalin Wadhwa, Jui Pradhan, Atharv Sonwane, Surya Prakash Sahu, Nagarajan Natarajan, Aditya Kanade, Suresh Parthasarathy, and Sriram Rajamani. 2023. Frustrated with Code Quality Issues? LLMs Can Help! *arXiv preprint arXiv:2309.12938* (2023).
- [135] Nalin Wadhwa, Jui Pradhan, Atharv Sonwane, Surya Prakash Sahu, Nagarajan Natarajan, Aditya Kanade, Suresh Parthasarathy, and Sriram Rajamani. 2024. CORE: Resolving Code Quality Issues using LLMs. *Proceedings of the ACM on Software Engineering* 1, FSE (2024), 789–811.
- [136] Han Wan, Hongzhen Luo, Mengying Li, and Xiaoyan Luo. 2024. Automated program repair for introductory programming assignments. *IEEE Transactions on Learning Technologies* 17 (2024), 1705–1720.
- [137] Shangwen Wang, Bo Lin, Liqian Chen, and Xiaoguang Mao. 2025. Divide-and-conquer: Automating code revisions via localization-and-revision. *ACM Transactions on Software Engineering and Methodology* 34, 3 (2025), 1–26.
- [138] Weishi Wang, Yue Wang, Shafiq Joty, and Steven CH Hoi. 2023. RAP-Gen: Retrieval-Augmented Patch Generation with CodeT5 for Automatic Program Repair. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. ACM, 146–158.
- [139] Xingyao Wang, Boxuan Li, Yufan Song, Frank F Xu, Xiangru Tang, Mingchen Zhuge, Jiayi Pan, Yueqi Song, Bowen Li, Jaskirat Singh, et al. 2024. OpenHands: An Open Platform for AI Software Developers as Generalist Agents. *arXiv preprint arXiv:2407.16741* (2024).
- [140] Yuxiang Wei, Olivier Duchenne, Jade Copet, Quentin Carbonneaux, Lingming Zhang, Daniel Fried, Gabriel Synnaeve, Rishabh Singh, and Sida I Wang. 2025. Swe-rl: Advancing llm reasoning via reinforcement learning on open software evolution. *arXiv preprint arXiv:2502.18449* (2025).
- [141] Yuxiang Wei, Chunqiu Steven Xia, and Lingming Zhang. 2023. Copiloting the Copilots: Fusing Large Language Models with Completion Engines for Automated Program Repair. In *Proceedings of the 31st ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 172–184.
- [142] Kyle Wong, Alfonso Amayuelas, Liangming Pan, and William Yang Wang. 2024. Investigating the transferability of code repair for low-resource programming languages. *arXiv preprint arXiv:2406.14867* (2024).
- [143] Fangzhou Wu, Qingzhao Zhang, Ati Priya Bajaj, Tiffany Bao, Ning Zhang, Ruoyu Wang, Chaowei Xiao, et al. 2023. Exploring the Limits of ChatGPT in Software Security Applications. *arXiv preprint arXiv:2312.05275* (2023).
- [144] Yi Wu, Nan Jiang, Hung Viet Pham, Thibaud Lutellier, Jordan Davis, Lin Tan, Petr Babkin, and Sameena Shah. 2023. How Effective Are Neural Networks for Fixing Security Vulnerabilities. In *Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis (ISSTA 2023)*. Association for Computing Machinery, 1282–1294.
- [145] Yonghao Wu, Zheng Li, Jie M Zhang, and Yong Liu. 2023. ConDefects: A New Dataset to Address the Data Leakage Concern for LLM-based Fault Localization and Program Repair. *arXiv preprint arXiv:2310.16253* (2023).
- [146] Marchel Christopher Wuisang, Marcel Kurniawan, Komang Andika Wira Santosa, Alexander Agung Santoso Gunawan, and Karen Etania Saputra. 2023. An Evaluation of the Effectiveness of OpenAI’s ChatGPT for Automated Python Program Bug Fixing Using QuixBugs. In *2023 International Seminar on Application for Technology of Information and Communication*. IEEE, 295–300.
- [147] Chunqiu Steven Xia, Yinlin Deng, Soren Dunn, and Lingming Zhang. 2025. Demystifying llm-based software engineering agents. *Proceedings of the ACM on Software Engineering* 2, FSE (2025), 801–824.
- [148] Chunqiu Steven Xia, Yifeng Ding, and Lingming Zhang. 2023. The Plastic Surgery Hypothesis in the Era of Large Language Models. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering*. IEEE, 522–534.
- [149] Chunqiu Steven Xia, Yuxiang Wei, and Lingming Zhang. 2023. Automated Program Repair in the Era of Large Pre-trained Language Models. In *2023 IEEE/ACM 45th International Conference on Software Engineering*. IEEE Computer Society, 1482–1494.
- [150] Chunqiu Steven Xia and Lingming Zhang. 2022. Less Training, More Repairing Please: Revisiting Automated Program Repair Via Zero-Shot Learning. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*. 959–971.
- [151] Chunqiu Steven Xia and Lingming Zhang. 2024. Automated Program Repair via Conversation: Fixing 162 out of 337 Bugs for \$0.42 Each using ChatGPT. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 819–831.
- [152] Jiahong Xiang, Xiaoyang Xu, Fanchu Kong, Mingyuan Wu, Zizheng Zhang, Haotian Zhang, and Yuqun Zhang. 2024. How far can we go with practical function-level program repair? *arXiv preprint arXiv:2404.12833* (2024).
- [153] Junjielong Xu, Ying Fu, Shin Hwei Tan, and Pinjia He. 2025. Aligning the Objective of LLM-Based Program Repair. In *2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)*. IEEE, 2548–2560.
- [154] Zhuolin Xu, Yuanzhang Lin, Qiushi Li, and Shin Hwei Tan. 2023. Guiding ChatGPT to Fix Web Ui Tests Via Explanation-Consistency Checking. *arXiv preprint arXiv:2312.05778* (2023).

- [155] Pengyu Xue, Linhao Wu, Zhen Yang, Zhongxing Yu, Zhi Jin, Ge Li, Yan Xiao, Shuo Liu, Xinyi Li, Hongyi Lin, et al. 2024. Exploring and Lifting the Robustness of LLM-powered Automated Program Repair with Metamorphic Testing. *arXiv preprint arXiv:2410.07516* (2024).
- [156] Weixiang Yan, Haitian Liu, Yunkun Wang, Yunzhe Li, Qian Chen, Wen Wang, Tingyu Lin, Weishan Zhao, Li Zhu, Shuiguang Deng, et al. 2023. Codescope: An Execution-Based Multilingual Multitask Multidimensional Benchmark for Evaluating LLMs on Code Understanding and Generation. *arXiv preprint arXiv:2311.08588* (2023).
- [157] Aidan ZH Yang, Sophia Kolak, Vincent J Hellendoorn, Ruben Martins, and Claire Le Goues. 2024. Revisiting unnaturalness for automated program repair in the era of large language models. *arXiv preprint arXiv:2404.15236* (2024).
- [158] Boyang Yang, Haoye Tian, Weiguo Pian, Haoran Yu, Haitao Wang, Jacques Klein, Tegawendé F Bissyandé, and Shunfu Jin. 2024. Cref: An llm-based conversational software repair framework for programming tutors. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 882–894.
- [159] Boyang Yang, Haoye Tian, Jiadong Ren, Hongyu Zhang, Jacques Klein, Tegawendé F Bissyandé, Claire Le Goues, and Shunfu Jin. 2024. MORepair: Teaching LLMs to Repair Code via Multi-Objective Fine-Tuning. *arXiv preprint arXiv:2404.12636* (2024).
- [160] John Yang, Carlos E Jimenez, Alexander Wettig, Kilian Lieret, Shunyu Yao, Karthik Narasimhan, and Ofir Press. 2024. SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering. *Advances in Neural Information Processing Systems* 37 (2024), 50528–50652.
- [161] John Yang, Carlos E Jimenez, Alex L Zhang, Kilian Lieret, Joyce Yang, Xindi Wu, Ori Press, Niklas Muennighoff, Gabriel Synnaeve, Karthik R Narasimhan, et al. 2024. SWE-bench multimodal: Do ai systems generalize to visual software domains? *arXiv preprint arXiv:2410.03859* (2024).
- [162] Xufeng Yao, Haoyang Li, Tsz Ho Chan, Wenyi Xiao, Mingxuan Yuan, Yu Huang, Lei Chen, and Bei Yu. 2024. HDLdebugger: Streamlining HDL debugging with Large Language Models. *arXiv preprint arXiv:2403.11671* (2024).
- [163] He Ye, Aidan ZH Yang, Chang Hu, Yanlin Wang, Tao Zhang, and Claire Le Goues. 2025. AdverIntent-Agent: Adversarial Reasoning for Repair Based on Inferred Program Intent. *Proceedings of the ACM on Software Engineering* 2, ISSTA (2025), 1398–1420.
- [164] Xin Yin, Chao Ni, Shaohua Wang, Zhenhao Li, Limin Zeng, and Xiaohu Yang. 2024. Thinkrepair: Self-directed automated program repair. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 1274–1286.
- [165] Zheng Yu, Ziyi Guo, Yuhang Wu, Jiahao Yu, Meng Xu, Dongliang Mu, Yan Chen, and Xinyu Xing. 2025. PATCHAGENT: A Practical Program Repair Agent Mimicking Human Expertise. In *Proceedings of the 34th USENIX Security Symposium (USENIX Security'25)*, Seattle, WA, USA.
- [166] Mingyue Yuan, Jieshan Chen, Zhenchang Xing, Aaron Quigley, Yuyu Luo, Tianqi Luo, Gelareh Mohammadi, Qinghua Lu, and Liming Zhu. 2024. Designrepair: Dual-stream design guideline-aware frontend repair with large language models. *arXiv preprint arXiv:2411.01606* (2024).
- [167] Wei Yuan, Quanjun Zhang, Tiek He, Chunrong Fang, Nguyen Quoc Viet Hung, Xiaodong Hao, and Hongzhi Yin. 2022. CIRCLE: Continual Repair across Programming Languages. In *Proceedings of the 31st ACM SIGSOFT International Symposium on Software Testing and Analysis*. ACM, 678–690.
- [168] Jialu Zhang, José Pablo Cambronero, Sumit Gulwani, Vu Le, Ruzica Piskac, Gustavo Soares, and Gust Verbruggen. 2024. PyDex: Repairing Bugs in Introductory Python Assignments using LLMs. *Proceedings of the ACM on Programming Languages* 8, OOPSLA1 (2024), 1100–1124.
- [169] Jiayi Zhang, Kai Huang, Jian Zhang, Yang Liu, and Chunyang Chen. 2025. Repair Ingredients Are All You Need: Improving Large Language Model-Based Program Repair via Repair Ingredients Search. *arXiv preprint arXiv:2506.23100* (2025).
- [170] Jian Zhang, Chong Wang, Anran Li, Wenhan Wang, Tianlin Li, and Yang Liu. 2024. Vuladvisor: Natural language suggestion generation for software vulnerability repair. In *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 1932–1944.
- [171] Kechi Zhang, Zhuo Li, Jia Li, Ge Li, and Zhi Jin. 2023. Self-Edit: Fault-Aware Code Editor for Code Generation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 769–787.
- [172] Lyuye Zhang, Kaixuan Li, Kairan Sun, Daoyuan Wu, Ye Liu, Haoye Tian, and Yang Liu. 2024. ACFIX: Guiding LLMs with Mined Common RBAC Practices for Context-Aware Repair of Access Control Vulnerabilities in Smart Contracts. *arXiv preprint arXiv:2403.06838* (2024).
- [173] Quanjun Zhang, Chunrong Fang, Weisong Sun, Yan Liu, Tiek He, Xiaodong Hao, and Zhenyu Chen. 2024. APPT: Boosting Automated Patch Correctness Prediction via Fine-Tuning Pre-Trained Models. *IEEE Transactions on Software Engineering* 50, 3 (2024), 474–494.



- [174] Quanjun Zhang, Chunrong Fang, Bowen Yu, Weisong Sun, Tongke Zhang, and Zhenyu Chen. 2023. Pre-Trained Model-Based Automated Software Vulnerability Repair: How Far are We? *IEEE Transactions on Dependable and Secure Computing* (2023). Early Access, DOI: [10.1109/TDSC.2023.3308897](https://doi.org/10.1109/TDSC.2023.3308897).
- [175] Quanjun Zhang, Chunrong Fang, Tongke Zhang, Bowen Yu, Weisong Sun, and Zhenyu Chen. 2023. GAMMA: Revisiting Template-based Automated Program Repair via Mask Prediction. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering*. IEEE, 535–547.
- [176] Quanjun Zhang, Tongke Zhang, Juan Zhai, Chunrong Fang, Bowen Yu, Weisong Sun, and Zhenyu Chen. 2023. A Critical Review of Large Language Model on Software Engineering: An Example from ChatGPT and Automated Program Repair. *arXiv preprint arXiv:2310.08879* (2023).
- [177] Ting Zhang, Ivana Clairine Irsan, Ferdian Thung, David Lo, Asankhaya Sharma, and Lingxiao Jiang. 2023. Evaluating Pre-Trained Language Models for Repairing Api Misuses. *arXiv preprint arXiv:2310.16390* (2023).
- [178] Yuwei Zhang, Zhi Jin, Ying Xing, and Ge Li. 2023. Steam: Simulating the Interactive Behavior of Programmers for Automatic Bug Fixing. *arXiv preprint arXiv:2308.14460* (2023).
- [179] Yuwei Zhang, Ge Li, Zhi Jin, and Ying Xing. 2023. Neural Program Repair with Program Dependence Analysis and Effective Filter Mechanism. *arXiv preprint arXiv:2305.09315* (2023).
- [180] Yuntong Zhang, Haifeng Ruan, Zhiyu Fan, and Abhik Roychoudhury. 2024. AutoCodeRover: Autonomous Program Improvement. In *Proceedings of the 33rd ACM SIGSOFT International Symposium on Software Testing and Analysis*. 1592–1604.
- [181] Jiuang Zhao, Donghao Yang, Li Zhang, Xiaoli Lian, Zitian Yang, and Fang Liu. 2024. Enhancing automated program repair with solution design. In *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering*. 1706–1718.
- [182] Qianhui Zhao, Fang Liu, Li Zhang, Yang Liu, Zhen Yan, Zhenghao Chen, Yufei Zhou, Jing Jiang, and Ge Li. 2024. Peer-aided Repairer: Empowering Large Language Models to Repair Advanced Student Assignments. *arXiv preprint arXiv:2404.01754* (2024).
- [183] Yuze Zhao, Zhenya Huang, Yixiao Ma, Rui Li, Kai Zhang, Hao Jiang, Qi Liu, Linbo Zhu, and Yu Su. 2024. RePair: Automated Program Repair with Process-based Feedback. In *Findings of the Association for Computational Linguistics: ACL 2024*. 16415–16429.
- [184] Zelin Zhao, Zhaogui Xu, Jialong Zhu, Peng Di, Yuan Yao, and Xiaoxing Ma. 2023. The Right Prompts for the Job: Repair Code-Review Defects with Large Language Model. *arXiv preprint arXiv:2312.17485* (2023).
- [185] Tianyu Zheng, Ge Zhang, Tianhao Shen, Xueling Liu, Bill Yuchen Lin, Jie Fu, Wenhui Chen, and Xiang Yue. 2024. OpenCodeInterpreter: Integrating Code Generation with Execution and Refinement. *arXiv preprint arXiv:2402.14658* (2024).
- [186] Li Zhong, Zilong Wang, and Jingbo Shang. 2024. LDB: A Large Language Model Debugger via Verifying Runtime Execution Step-by-step. *arXiv preprint arXiv:2402.16906* (2024).
- [187] Xin Zhou, Kisub Kim, Bowen Xu, DongGyun Han, and David Lo. 2024. Out of Sight, Out of Mind: Better Automatic Vulnerability Repair by Broadening Input Ranges and Sources. In *2024 IEEE/ACM 46th International Conference on Software Engineering*. IEEE Computer Society, 872–872.
- [188] Xin Zhou, Bowen Xu, Kisub Kim, Dong Gyun Han, Hung Huu Nguyen, Thanh Le-Cong, Junda He, Bach Le, and David Lo. 2024. Leveraging Large Language Model for Automatic Patch Correctness Assessment. *IEEE Transactions on Software Engineering* 50, 11 (2024), 2865–2883.
- [189] Armin Zirak and Hadi Hemmati. 2024. Improving Automated Program Repair with Domain Adaptation. *ACM Transactions on Software Engineering and Methodology* 33, 3 (2024), 1–43.