

JUNHA PARK

junha.p@snu.ac.kr | junha.me | github.com/Oxjunha | linkedin.com/in/Oxjunha | Seoul, South Korea

Experience

Independent Benchmark Engineering Independent Contractor	May 2026 – Present
Review coding-agent benchmark tasks for private evaluation suites.	
FastRoll Repository Founder & Software Engineer	Jul 2024 – Present (Full-time Apr 2025 – Feb 2026)
- Implemented a Rust client for the Join-Accumulate Machine (JAM) protocol directly from the Gray Paper formal specification, without a reference implementation, spanning state transition, storage, erasure coding, Merkle structures, and a RISC-V VM interpreter. - Passed JAM Prize Milestone 1 conformance fuzzing by importing 1.3 million generated blocks and producing valid post-states, with the result independently attested by Parity Technologies.	
Aptos Labs Software Engineer	Jan 2024 – Apr 2025
Shipped features across Aptos's Rust/Move framework, CLI, and indexer stack, including chunked publishing of large smart-contract packages, account restoration for rotated and multi-key accounts, and gas optimization for the BCS deserializer.	
Rise Finance Co-founder & Software Engineer	Jan 2023 – Jan 2024
Built the smart-contract trading engine for a pre-launch perpetual futures exchange: price oracle, settlement, and liquidation.	
LRHR (FAPIS) Software Engineer	Feb 2022 – Aug 2022
Built the initial backend, deployment, payment, and settlement services for a luxury-repair marketplace as its first engineer.	

Open-Source Work

Terminal-Bench 3 & 4 Merged Task Benchmark Task Author	May 2026 – Present
- Designed and upstreamed rs-archive-clone, a task requiring a clean-room reimplementaion of a layered archive/recovery tool. Claude Fable 5 and GPT-5.6 Sol each passed 1 of 5 trials (20%); each of the other nine models passed 0 of 5. - Audited agent trajectories and found models copying the reference binary and obfuscating its path to evade detection. A follow-up PR (open) replaces the task's heuristic reward-hacking detector with structural isolation, moving the reference into a probe-only sidecar. - Ported the task to Inspect.	
Gray Paper Repository Authored Commits Editorial Board Member	Jul 2025 – Present
Review protocol upgrades, resolve ambiguities with protocol authors and implementers, and author specification corrections across PVM semantics, state-transition logic, erasure coding, and serialization.	

Projects

RISC-V VM Speedrun Benchmark Repository Benchmark Task Author	Jul 2026 – Present
- Designed a single-task long-horizon benchmark that measures the architectural judgment of coding agents rather than their raw task-solving capability. Agents receive one of six RISC-V VM implementations of differing maturity and must make it faster. - Ran 30 preliminary 6-hour trials on gpt-5.6-sol and found path dependence. Every run starting from a mature Rust JIT underperformed every run starting from a naive Python interpreter. - Isolated each trial, re-baselining to hold baseline variation under 0.4%.	
Darc Repository Software Engineer	Apr 2026 – Jul 2026
Built a Rust CLI that incrementally normalizes and indexes Codex and Claude Code rollouts into SQLite, supporting structured queries over sessions, turns, tool calls, and touched files with exact evidence retrieval for analyzing agent behavior.	

Skills

Languages	Rust, Python, TypeScript, SQL	Evaluation Tooling	Inspect, Harbor (task authoring & analysis)
------------------	-------------------------------	---------------------------	---

Education

Seoul National University B.S. in Computer Science and Engineering	Expected Aug 2027
--	-------------------