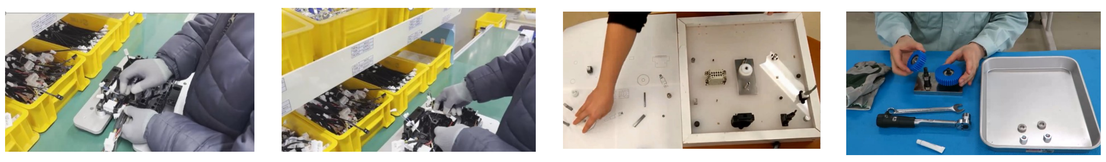
Hybrid Agentic AI-FSM Framework for Instruction-based Industrial Manipulation Tasks

Sungmoon Joo*, Ikjune Kim
Korea Atomic Energy Research Institute

Motivation & Challenges

• In small- to medium-sized factories with flexible production lines, robotic systems must handle frequently changing assembly tasks without requiring time-consuming reprogramming.

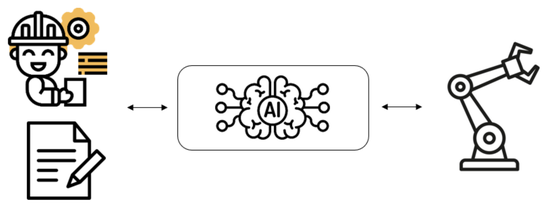


• **Industrial Needs:** High-reliability, high-risk, and safety-critical environments—such as manufacturing—demand robot stability that directly impacts productivity.

• **Gap in Automation:** While industrial tasks are typically documented in natural language (SOPs), converting these instructions into executable robot programs currently requires intensive manual effort.

• **Limitations of Current AI:** State-of-the-art Vision-Language-Action (VLA) models are "black-box" systems prone to hallucinations and non-determinism, making them unverifiable by industrial safety standards.

• **The Dilemma:** A fundamental conflict exists between the flexible reasoning of LLMs and the deterministic reliability required for industrial control.



Proposed Solutions: Hybrid Agentic AI-FSM Framework

Key Idea

• A complex task is decomposed into simple unit tasks.
• A language model, restricted to offline planning, generates a unit-task-level task plan.
• A verified FSM-style execution engine manages the deterministic execution of the task plan.

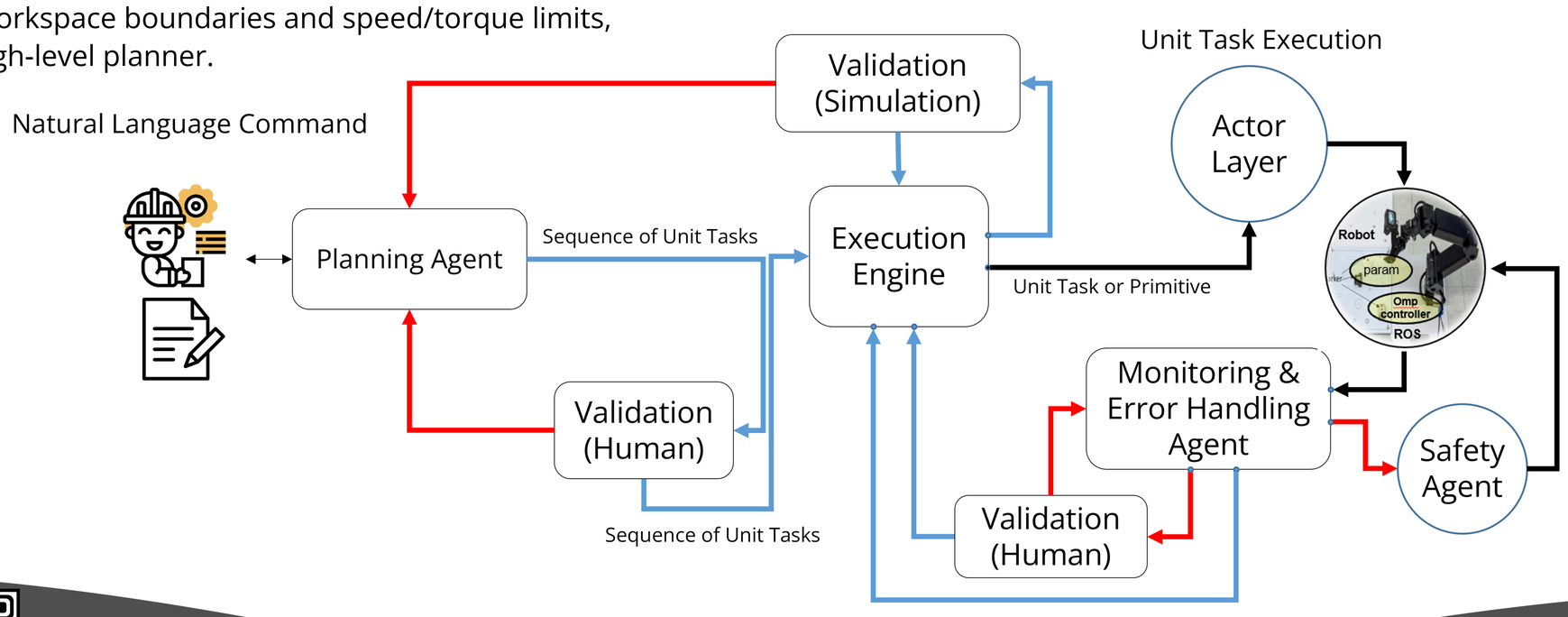
Key Elements

• Hierarchical Task Planning: Instructions are decomposed into a three-level hierarchy—Composite Tasks, Unit Tasks, and Primitive Actions—ensuring semantic grounding and valid execution.

• Multi-stage Validation Pipeline: Logical and physical validity is verified through schema checks, rule-based constraints, and simulation-based dry-runs before any real-world deployment.

• RAG-Enhanced Exception Handling: Upon detecting runtime anomalies, a RAG-based agent retrieves relevant manuals and proposes recovery actions. The proposed recovery plan must be approved by a human operator (Human-in-the-Loop) before task execution resumes.

• Independent Safety Agent: A dedicated physical filter enforces real-time constraints, such as workspace boundaries and speed/torque limits, independent of the high-level planner.



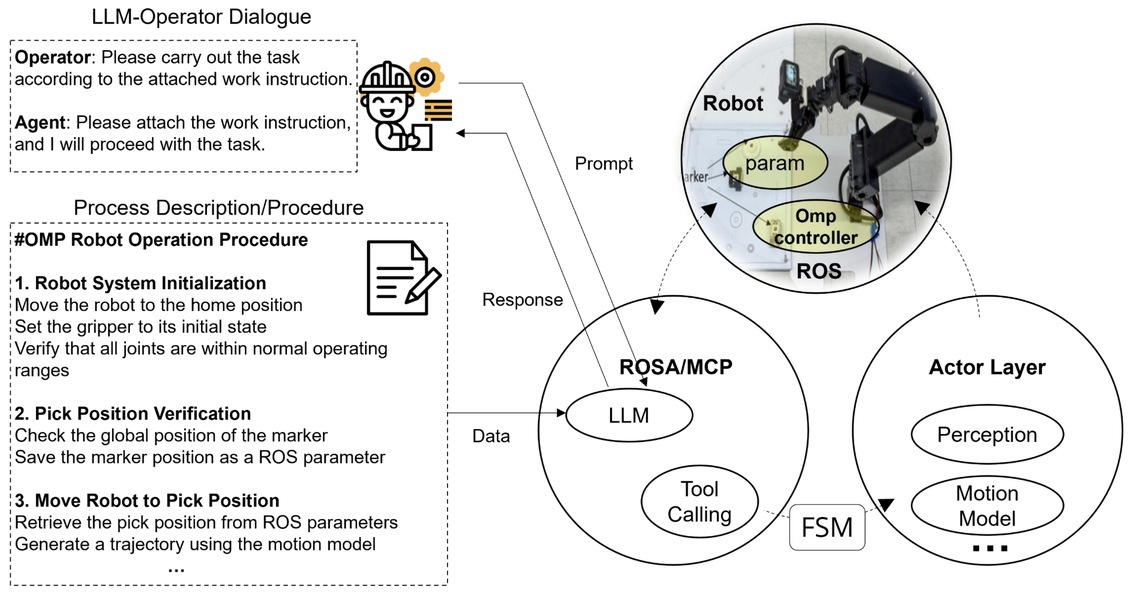
Implementation

Agentic AI Framework for Robot Control

• ROSA: ROSA is an AI-powered assistant for ROS1 and ROS2 systems. Built on the Langchain framework, ROSA enables operators to interact with robots using natural language (https://github.com/nasa-jpl/rosa)

• ROS-MCP-SERVER: ROS-MCP-SERVER connects large language models (such as Claude, GPT, and Gemini) to robots, enabling bidirectional communication with no changes to existing robot source code (https://github.com/robotmcp/ros-mcp-server)

Component	Implementation Tool	Role & Responsibility
Planner Agent	**ROSA** (LangChain-based)	High-level reasoning and **Unit Task** composition.
Monitoring & EH Agent	**ROSA**	Real-time log/event auditing and autonomous recovery.
High-level Safety Agent	**ROSA**	Strategic safety validation (Logical Fail-safe).
Context Gateway	**ROS-MCP-SERVER**	Exposes ROS context to external LLM clients (Claude, Cursor).
Execution Engine (Outside Agent)	**BehaviorTree** or **State Machine**	Deterministic execution of the Unit Task sequence.
Low-level Safety Agent (Outside Agent)	**Native C++ / ROS Node**	Real-time collision avoidance and hardware protection.
Actor Layer (Outside Agent)	**PyTorch** (ROS Inference Node)	Execution of **Imitation-Learned (IL)** policies.



Summary & Conclusion

• The proposed framework implements a multi-layered protection strategy (aka "Defense-in-Depth“ in nuclear industry), integrating logical guardrails with physical safety controllers to mitigate both reasoning errors and physical risks.

• This hybrid approach overcomes the uncertainty of end-to-end models, offering a practical, explainable, and safe solution for natural-language-driven industrial automation.

Hybrid Agentic AI-FSM Framework

< Contact >
Sungmoon Joo, PhD
Principal Researcher
Email: smjoo@kaeri.re.kr
Website: https://pairlab.net

< Acknowledgment >
This work was supported by the Ministry of Trade, Industry and Energy, Republic of Korea, under Grant No. RS-2024-00416440