

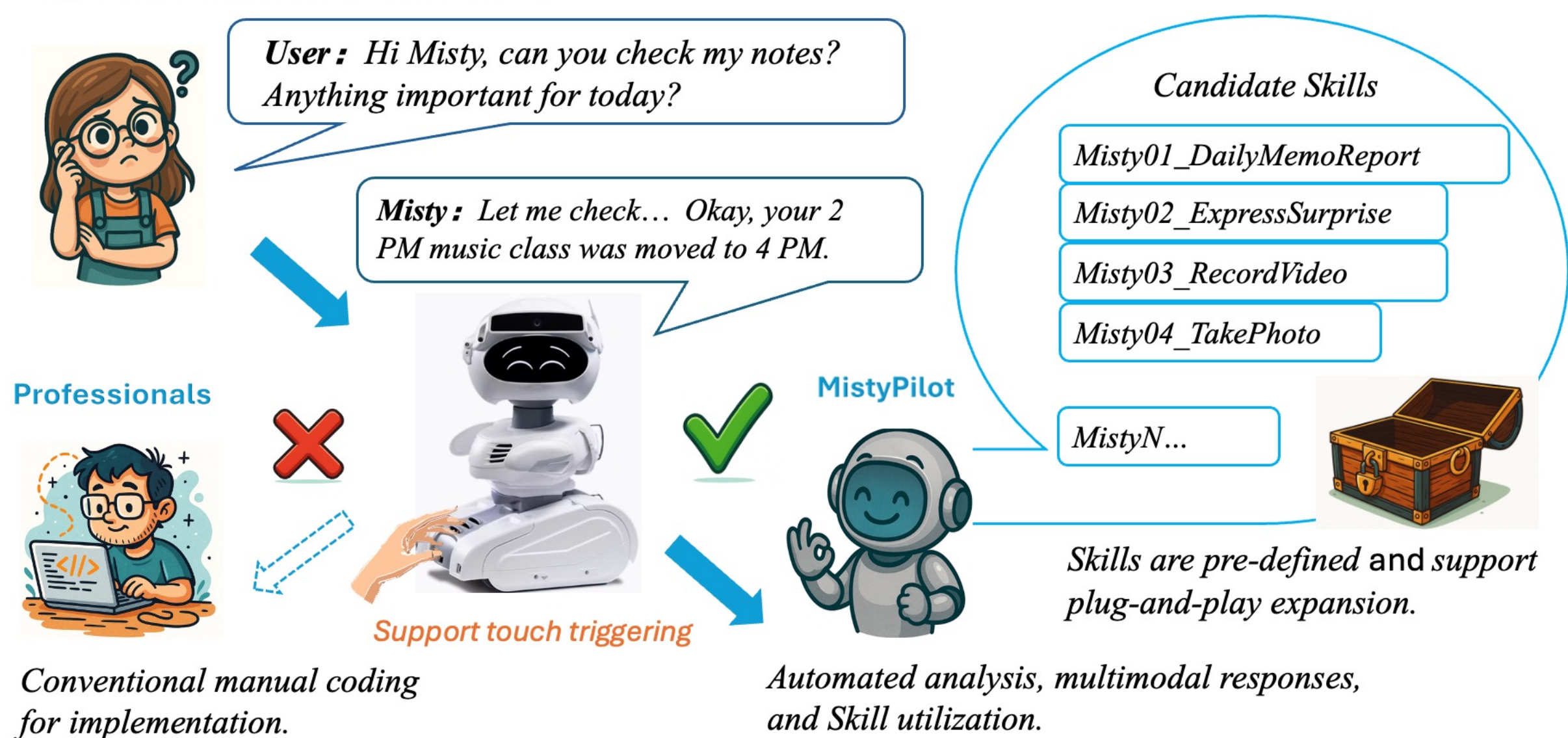
Motivation

How can we translate user requirements into executable robot behaviors?

For a specific task, users still need to:

- **Program** robot actions.
- **Define** triggers and conditions.
- **Design** responses and interaction flows.

High-level Human Instructions



From Skill Generation to Skill Orchestration

AutoMisty

PRIOR WORK · IROS 2025

Generate skills (control code) for open-ended tasks

- Generate robot skills from natural-language requests.
- Execute and refine the skills to complete the task.

MistyPilot

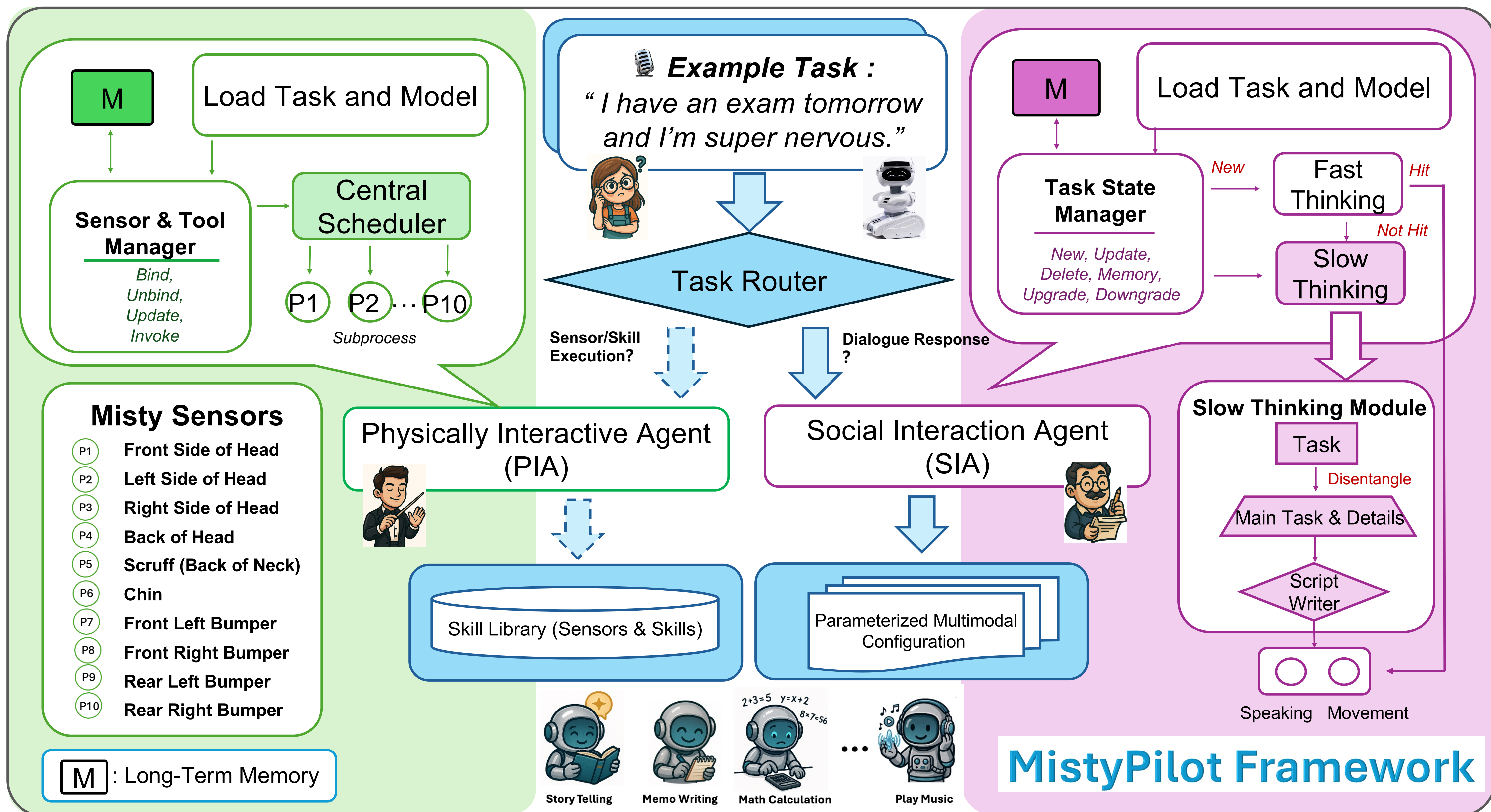
THIS WORK · ECCV 2026 ACVR

Orchestrate skills during interaction

- Select and configure skills based on user requests.
- Use sensor bindings and dialogue state to guide execution.

MistyPilot Framework

MistyPilot's Multi-Agent Architecture



Physically Interactive Agent (PIA)

- Select and configure robot skills (control code) from user instructions.
- Execute once or maintain persistent sensor-skill bindings.
- Select new skills by description and configure them in the system.

Social Interaction Agent (SIA)

- Coordinate emotional tone, intonation, speaking rate, motion, and expressions.
- Reuse results users choose to save; otherwise generate a new response.
- Maintain explicit task state and selectively update details across turns.

Evaluation

Technical Evaluation

Consistently high accuracy

With GPT-5-mini as its backbone, MistyPilot achieved over **95%** mean accuracy across all evaluated components.

Faster responses through result reuse

Core-task retrieval reuses user-saved results, reducing mean response time from **5.09 s** to **2.26 s**—a **55.5%** reduction.

Robust skill-library expansion

No degradation in skill-selection performance was observed as the library expanded from **30** to **100** skills.

User Study

12 users · 30+ min each · 5-item Likert survey

Dimension	Mean	SD
Interaction naturalness	4.75	0.45
Comprehension accuracy	4.42	0.67
Emotional expressiveness	4.75	0.45
System responsiveness	4.67	0.49
User satisfaction	4.75	0.45

All five mean ratings exceeded 4.4/5.



This material is based upon work supported under the AI Research Institutes program by the U.S. National Science Foundation and the Institute of Education Sciences, U.S. Department of Education, through Award 2229873—National AI Institute for Exceptional Education. Any opinions, findings and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, the Institute of Education Sciences, or the U.S. Department of Education.

