Mechanics-Aware LLM Orchestrator for Human–Robot Bimanual Co-Manipulation

Saien Deng, David Scott Lewis
AIXC, Spain

Introduction and motivation

robot and human coordination becomes increasing important these days, especially under the scenario of co-manipulation of objects. Under such circumstances, internal force management becomes a challenge as force imbalances would cause safety issues and compliance was required. Though fundamental control strategies (Hybrid position/force control, impedance control and operational space control) provided insights in addressing such issue, question still arises on how to adopt the optimal strategy (stiffness, damping, coordination modes) between these in complexed task Situations.

To address such issue, we proposes the idea of Mechanics-Aware LLM Orchestrator, a system that contains a pretrained large language model tunes low level controller and end effectors through API keys based on human instructions given. This system integrates a state-of-the-art bimanual controller with a LLM oracle (Fig. 1)

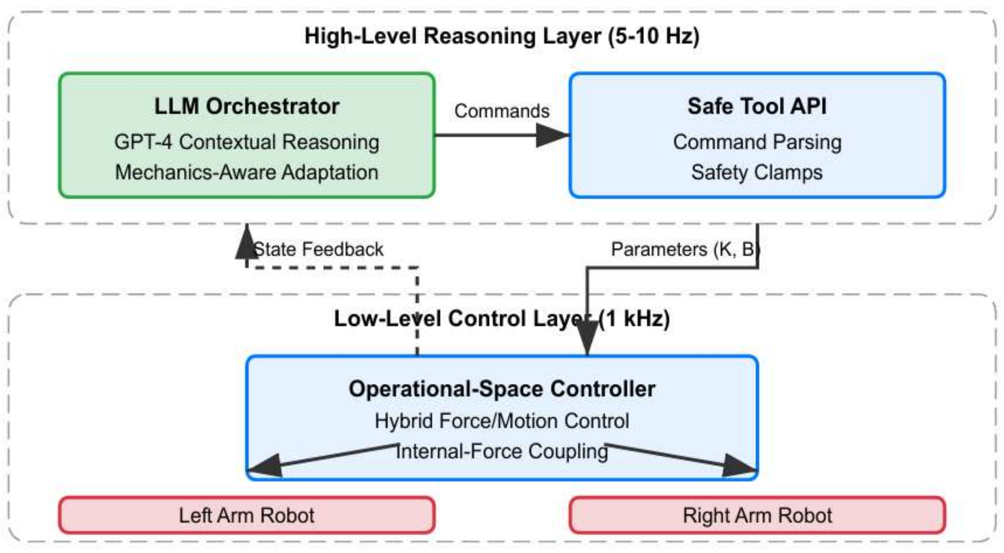


figure 1

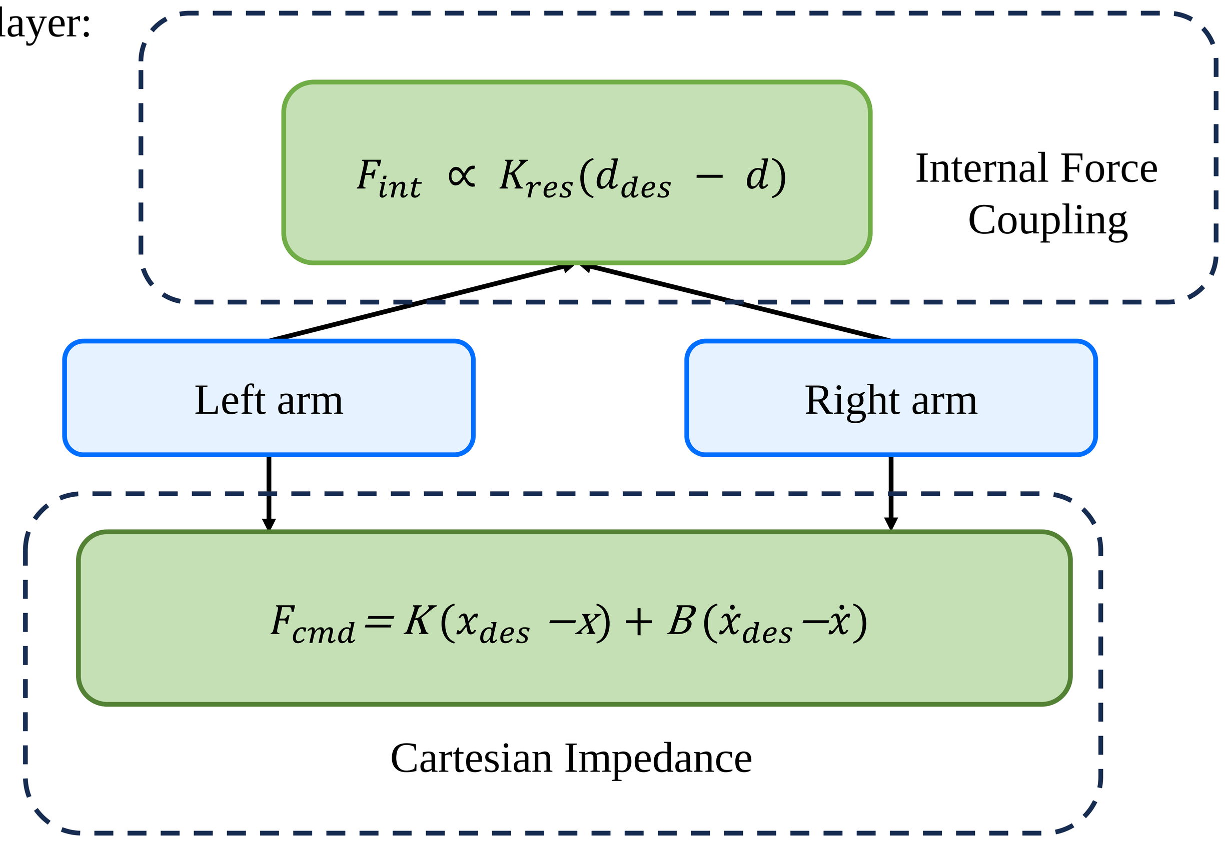
.

Architecture

Our framework is separated into 2 layers:

1. real-time Hybrid Operational-Space Controller that drives the robot's arms
2. the LLM Orchestrator, which monitors the task and issues high-level adjustments via a textual command interface.

First layer:



Second layer: LLM Orchestrator and API Interface

1. State observation encoding:
   example: `State:`
   ```
   Task: Carry table to corner.
   Phase: Walking, human leading.
   Left/Right Load: 12 N / 8 N (Target 10 N).
   ```
2. Mechanics-Aware Tool API:
   example: `choose_mode(seq_or_sim)//switch between sequential or simutanious action`
   `replan_grasps(strategy)// switch strategy`
3. Chain-of-Thought for Stability:
   example :`<robot>: human speeding up > adjust to low stiffness`

System Execution and Safety:

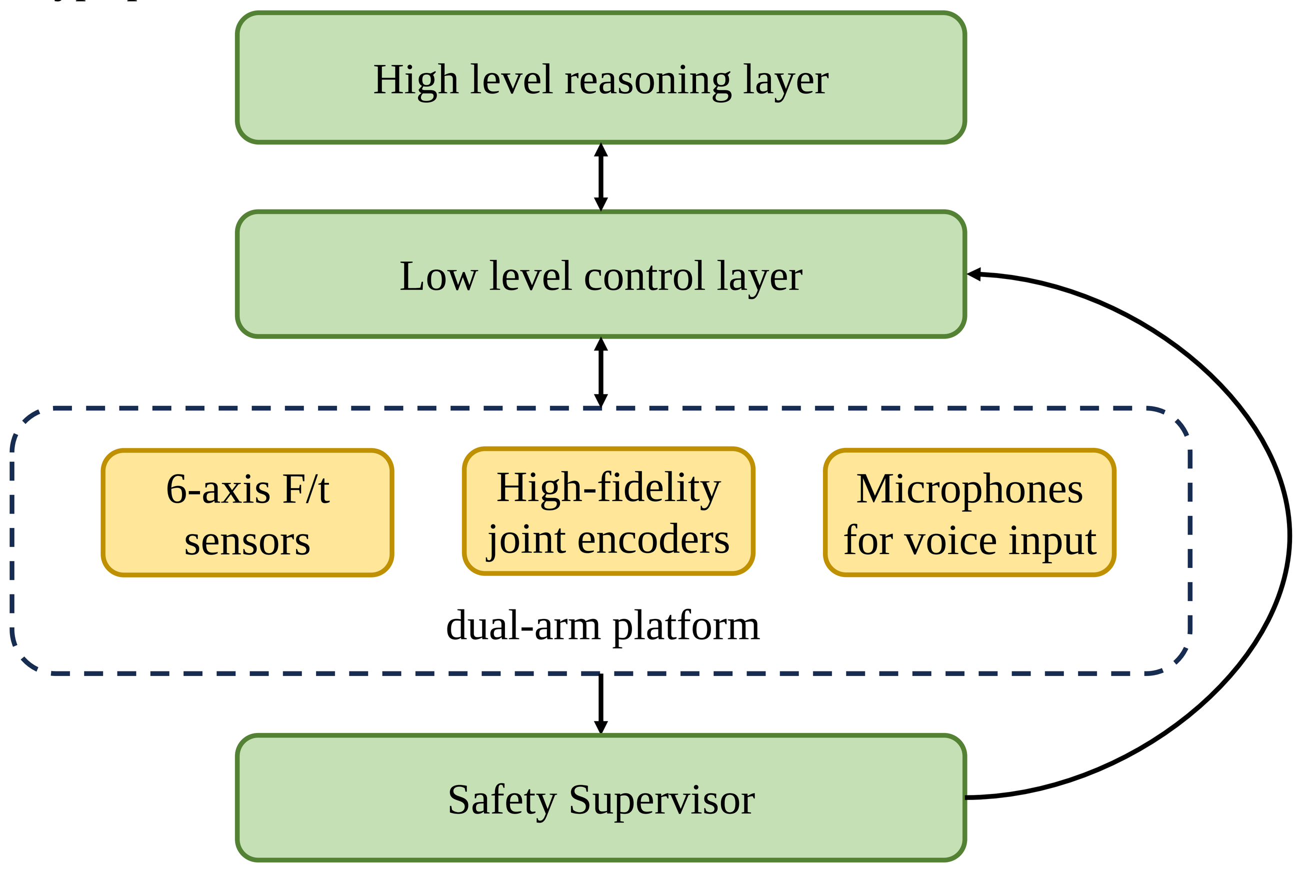
the control loop delivers message between different counterparts. To ensure safety. Limits are set on significant change. API interpreters and low-level controls will also stop LLM value inputs when it exceeds limit.

Prompt design

1. Role and High-Level Goal: introduced long time objective on safety and collaboration
   example: `you are a robotic assistant embodied in a dual-arm body, your task is to assist human.`
2. Tool description and hints: API document alongside with instructions
   example: `set_impendance(arm, Level)// low makes the arm compliant (preferred if human lead), high makes arm stiff (preferred for support)`
3. few shot examples: samples of observation, chain of thought, and tool use provided
   example: Jar lid stuck > need bigger torque with caution in speed > `choose_mode(seq) set_impedance(high)`
4. Safety reminders and constraints: a list of caution provided to the model
   example: `Never exceed speed limit of 5m/s.`

Prototype plan and ablation studies

Prototype plan:



Evaluation and ablation study:

scenarios: 1. collective dual hand table carry. Test goal: load sharing and responsiveness to human based challenge
2. Asymmetric Bimanual Task (Jar Twist). Test goal: internal force balance

Baselines

B1: Fixed-Gain Impedance: Constant medium stiffness. No adaptation.
B2: Independent Arms with only force impendence
B3: LLM with No Mechanics API
Proposed: LLM Orchestrator

Ablation:

TABLE I
HYPOTHESIZED ABLATION STUDY RESULTS (CO-MANIPULATION)

Method	Peak	Int. Var.	Fluency	Adapt.
Fixed-Gain (B1)	High	Med.	Med.	None
Indep. Arms (B2)	V. High	High	Low	Low
LLM w/o Mech. (B3)	High	Med.	Med.	High-lvl
LLM Orch. (Full)	**Low**	**Low**	**High**	**Cont.**

conclusions and future work

Future work: 1. improve safety measures with a safety supervisor or verifiable feedback
2. adapt event triggered approach and achieve latency proposed,
3. integrate multimodal inputs by integrating frameworks like VLA
4. integrate task planner for long-horizontal tasks.

Conclusion: we provided a conceptual framework that combines Large language models with fine tuned control of robots. By grating LLM the access to robot Control loops, we enabled a form of intelligent, instruction aware assistant for human life.