

Dynamic Penetrative Trajectory Adaptation ??

Zhifeng Han. Author ?? Clarire Walton. Author, Jr. ??

* University of Texas at San Antonio, San Antonio, TX 78249 USA
(e-mail: zhifeng.han@utsa.edu).

** University of Texas at San Antonio, San Antonio, TX 78249 USA
(e-mail: claire.walton@utsa.edu).

Abstract: DPTA is a conceptual framework where the LLM's world knowledge is leveraged to analyze low-level sensor feedback, enabling the dynamic selection and refinement of complex, pre-learned robot trajectories to achieve high dexterity and resilience in autonomous manipulation.

Keywords: LLM, robotic, Reinforcement Learning.

1. INTRODUCTION

interactive robots mission is a complex mission it need to navigate in complex and dynamic environment also need to fully flexible to planing the sequentially sub tasks. In order to achieve the goal we demo a method by utilize LLMs(large language model) and Reinforcement Learning. Our idea focuses on using an LLM's capacity for precise, short-horizon physical reasoning to dynamically select or modify the parameters of complex, learned robotic trajectories, effectively allowing the robot to autonomously execute tasks that previously required initial human intervention. The conceptual framework of **Dynamic Penetrative Trajectory Adaptation (DPTA)** represents a synthesis of three complementary AI paradigms. First, *Penetrative AI* enables contextual and semantic understanding by leveraging large language model-based sensor comprehension for high-level intent extraction and environmental interpretation. Second, the combination of *Dynamic Movement Primitives (DMP)* and *Human-Robot Collaboration (HRC)* contributes to adaptable trajectory generation through shared autonomy, interactive learning, and human demonstration. Third, *Robust Reinforcement Learning (RL)* performs policy optimization and closed-loop execution under uncertainty, ensuring real-time correction and robust action adaptation. Together, these three paradigms form the foundation of the DPTA framework, enabling semantic reasoning, dynamic motion modification, and reliable trajectory shaping in complex robotic environments.

2. BACKGROUND

The complex nature of interactive robot missions, particularly those involving high-level language models (LLMs) or sophisticated reinforcement learning (RL), presents numerous challenges across planning, learning, perception, and execution.

* Sponsor and financial support acknowledgment goes here. Paper titles should be written in uppercase and lowercase letters, not all uppercase.

2.1 Main challenges

Limitations in High-Level Planning and Trajectory Execution (LLMs): When LLMs are used for autonomous manipulation, they encounter issues related to generating and executing physical motions: Inability to Handle Complex Trajectories(?) (Feasibility Issues): The conventional approach of using LLMs to generate code for robot motion falters when dealing with complex trajectories. Tasks that require intricate trajectory planning and reasoning over environments, such as opening an oven door featuring a horizontal axis design or opening a cabinet with a press-pull structure, may be deemed infeasible when relying only on the basic motion library generated by the LLM(?). Fragility of Prompt Design: The current design paradigm for using LLMs as controllers is fragile, meaning even minor alterations in the prompt can dramatically affect the performance.? Designing a reliable prompt for robotic tasks is not yet well understood.

Executability Anomalies: Although generally high, the code generated by the LLM can occasionally generate sub-tasks without assigning corresponding motion functions, resulting in non-encodable and non-executable responses(?).

Issues Related to Perception and Grounding:

Successfully linking high-level instructions to the physical world introduces multiple errors: Environmental Perception Errors and Error Accumulation: Real-world task success rates decrease due to error accumulation across sequential sub-tasks(?). Errors in environmental perception stem from inaccuracies in object detection models (like YOLOv5), such as bounding box inaccuracies, leading to slightly variable coordinates for target objects.? These discrepancies can cause the errors to exceed the necessary margins for precise manipulation (e.g., placing an apple into an oven with minimal clearance)(?) Sensor Data Processing Limitations: LLMs, when used in a "penetrative" way to analyze digitized sensor signals (like sequences of ECG digits), exhibit lower efficiency in processing ex-

tensive sequences of digital data compared to traditional methods(?). The hallucination rates and Mean Absolute Errors (MAEs) for some LLMs escalate with the increase in window size of the input data, suggesting an inherent limitation in processing extensive lengths of digitized sequences(?). Susceptibility to Deployment Noise: Policies trained in simulation, even when using modern techniques, may not be robust to real-world noise. For instance, a policy trained in simulation for pick-and-place was not robust to small errors in box position estimation (e.g., errors with a standard deviation of 1cm) when deployed on a physical robot(?).

Challenges in Low-Level Control:

Controllability and Security Risks: Since LLM responses are probabilistic, there is no guarantee that the swarm will behave as intended. This also introduces new security vulnerabilities, as it needs to be studied if users or even other robots can reprogram robots through prompt injection attacks or if a malicious agent could send misleading information (Byzantine robot detection)?.

2.2 Overview: Dynamic Penetrative Trajectory Adaptation (DPTA)

In current LLM-based manipulation, environmental information is primarily derived visually (e.g., YOLOv5 for object position)?). However, fine-grained tasks often depend on non-visual physical feedback (e.g., force or torque required to open a tight hinge). In DPTA, the Penetrative AI(?) paradigm is employed to process digitized sensor signals** from the robot’s end effector (e.g., force/torque sensors, joint current feedback)

- LLM’s task is not to determine a broad state (like “indoors/outdoors”), but to execute a real-time, micro-level physical classification of the object/environment state during the initial phase of interaction (e.g., the first 100 milliseconds of grasping a handle or pushing a button)?.
- Prompt Design: Our prompt utilizes the procedural guidance and fuzzy logic methods demonstrated in the heart rate detection task.(?) It contains a short sequence of raw numerical feedback (avoiding the token limit constraint associated with long digitized sequences) and instructs the LLM to classify the physical anomaly based on relative changes in the sequence.
- Example Output: Based on the input torque sequence, the LLM reasoning(?) determines the precise physical condition, such as “Horizontal-Axis Oven Door, stiff hinge” or “Press-Pull Cabinet, high friction.”
- Dynamic Controller Role: The LLM functions as a **dynamic feedback controller**, selecting between:

2.3 Literature Review

The proposed Dynamic Penetrative Trajectory Adaptation (DPTA) framework differentiates itself from existing research in LLM-controlled robotics by addressing limitations in physical grounding, low-level trajectory execution, and robust policy learning.

- (1) **High-Level Planning with Imperfect Grounding** Existing works leverage large language models (LLMs) as **zero-shot planners** to decompose high-level instructions (e.g., “make breakfast”) into a sequence of actionable steps [?]. Methods like those involving **Semantic Translation** improve executability by mapping LLM-generated phrases to the most semantically similar admissible action in a predefined set. **SayCan** grounds LLM output by weighting actions based on learned **skill affordances** (value functions), ensuring the proposed step is feasible in the current state [?].

Limitation: Reliance on Narrow APIs and Lack of Deep Physical Context

Despite these efforts, LLM-generated plans are still frequently not executable in interactive environments (?) and struggle with mid-level grounding, often missing necessary common-sense actions [?]. Critically, most systems rely on predetermined **vision APIs** (like object detectors) to describe the scene [?]. This focus on visual data ignores crucial non-visual **physical feedback** (e.g., force or torque) needed for fine-grained tasks [?], leaving the LLM policies restricted by what the perception APIs can describe [?].

- (2) **Code Generation and Trajectory Management**

Recent research utilizes LLMs to directly generate **policy code (Code as Policies)**, allowing the model to compose perception-to-control logic and reference external libraries (like NumPy) for complex behaviors such as spatial-geometric reasoning [?]. Furthermore, LLMs have been explored as **low-level feedback policies** for dynamic systems like robot walking by outputting target joint positions directly from historical input-output sequences [?].

- (3) **Closed-Loop Adaptation and Policy Learning**

Systems such as **Inner Monologue** and **Socratic Models** integrate closed-loop feedback by feeding natural language observations (e.g., success detection or scene description updates) back into the LLM prompt, enabling the LLM planner to reason over outcomes and dynamically re-plan actions [huang2023inner]. In the domain of reinforcement learning (RL), solving robotic tasks often faces the critical challenge of **sparse and binary rewards** [?].

- (4) **Feedback Limitations and Learning Inefficiency**

The effectiveness of closed-loop reasoning is **bottlenecked** by the capabilities of the low-level control policies and the fidelity of the language description provided by the perception system [?, ?]. Moreover, traditional RL systems often fail in large state spaces when faced with sparse rewards, often necessitating tedious and domain-specific **reward function engineering** [?].

3. ARCHITECTURE

Penetrative Trajectory Adaptation (DPTA) architecture is a composite framework that integrates Large Language Models (LLMs) with specific sensing and control methodologies to address the limitations of grounding and trajectory generation in robotics. And we propose in this diagram our primary layers:

$$\mathcal{A}_{DPTA} = \{\mathcal{L}_{plan}, \mathcal{L}_{perc}, \mathcal{L}_{exec}, \mathcal{L}_{learn}\}.$$

The Dynamic Penetrative Trajectory Adaptation (DPTA) framework is a hierarchical robotic architecture that integrates large-scale language-based reasoning, physically grounded sensory interpretation, adaptive trajectory learning, and reinforcement learning. The system enables a robot to interpret natural language instructions, perceive through multimodal signals, and adapt its motion in real-world conditions.

3.1 High-Level Reasoning and Planning Layer

At the highest level, a Large Language Model (LLM) serves as the central cognitive planner. Given a natural-language instruction (e.g., “warm up my lunch”), the LLM performs long-horizon task decomposition and generates a structured sequence of sub-tasks. Each sub-task is mapped to motion primitives selected from a predefined Basic Motion Library. This hierarchical formulation allows complex goals to be broken into intermediate actions, which are subsequently instantiated via control functions (e.g., `move_to_pose`, `align_gripper`, `apply_force`). This layer supports generalized planning without task-specific programming.

3.2 Penetrative Perception Layer

Unlike vision-centric pipelines that rely solely on symbolic or labeled information, DPTA leverages Penetrative AI for physically grounded perception. In this layer, raw sensor signals (IMU sequences, tactile arrays, periodic biosignals, etc.) are directly fed into the LLM rather than intermediate or symbolic features. Using embedded world knowledge and numerical reasoning, the LLM interprets these signals to infer latent physical properties (e.g., material stiffness, friction, compliance, or slippage). This enables the architecture to detect micro-level physical properties that are difficult or impossible to capture using visual sensing alone.

3.3 Dynamic Execution Layer

For tasks that exceed the capabilities of predefined motion primitives, DPTA incorporates Dynamic Movement Primitives (DMPs) in combination with Human-Robot Collaboration (HRC). The system maintains a DMP Library storing parameterized, high-dimensional manipulation skills. When the LLM determines that a trajectory cannot be synthesized from the Basic Motion Library, the system allows a human operator to intervene via teleoperation. The demonstrated trajectory is encoded as a DMP using Gaussian basis functions and a corresponding weight vector. The resulting representation is added to the library for future reuse, enabling continual expansion of manipulation skills and reusable motor behaviors.

3.4 Robust Learning Layer

To ensure continual improvement in manipulation performance, DPTA integrates Hindsight Experience Replay (HER). Many manipulation tasks yield sparse binary rewards, which impede learning in standard reinforcement

learning formulations. HER alleviates this challenge by relabeling unsuccessful episodes as successful transitions toward alternative goals. A replay buffer stores trajectories and reinterprets the achieved end-state as the desired goal. This implicit curriculum accelerates policy learning and improves robustness across failures, allowing the system to autonomously improve over time with minimal human supervision.

$$DPTA = \mathcal{F}(\mathcal{P}_{PenAI}, \mathcal{T}_{DMP/HRC}, \mathcal{R}_{RRRL})$$

$$DPTA = \underbrace{\mathcal{P}_{PenAI}(s)}_{\text{Semantic Interpretation}} \Rightarrow \underbrace{\mathcal{T}_{DMP/HRC}(\cdot)}_{\text{Trajectory Generation}} \Rightarrow \underbrace{\mathcal{R}_{RRRL}}_{\text{Policy Execution}}$$

4. IMPLEMENTATION DETAILS

DPTA Validation: Adaptive Cabinet Opening under Variable Friction

5. EXPERIMENTAL VALIDATION OF DPTA ON A CABINET-OPENING TASK

5.1 Objective

The experiment aims to demonstrate that an LLM-based system can adapt a learned Dynamic Movement Primitive (DMP) to varying physical constraints (e.g., a stuck or heavy drawer) by interpreting penetrative, non-visual sensor data (joint torques and forces). Visual-only planners are expected to fail under these conditions, whereas the DPTA framework leverages this additional sensory information for robust adaptation.

5.2 Apparatus and Environment

- **Robot:** UFactory Lite 6, simulated in NVIDIA Isaac Sim and validated on the real robot for demonstration.
- **Control Software:** UFactory Studio for Human-Robot Collaboration (HRC) and trajectory recording.
- **Simulation Engine:** NVIDIA Isaac Sim / Isaac Gym for physics simulation and massive parallel training.
- **Task:** “Open the Cabinet”.
- **Variable Parameters:** Friction coefficient of the cabinet slide and mass of the cabinet door.

5.3 Methodology

Phase 1: Trajectory Acquisition via HRC **Goal:** Acquire the base DMP.

Procedure:

- (1) Set the UFactory Lite 6 to Manual (Teach) Mode with gravity compensation.
- (2) A human operator guides the robot to grasp and open the cabinet, producing the “Golden Trajectory”.
- (3) Record joint positions and velocities at 250 Hz using UFactory Studio.
- (4) Encode the recorded trajectory into a DMP using Gaussian basis functions to create a flexible, parameterized motion primitive. This constitutes the *Base Skill* stored in the library.

Phase 2: Penetrative Simulation Environment Goal:
Simulate invisible physical constraints and expose penetrative sensor signals.

Setup:

- (1) Import the UFactory Lite 6 URDF into Isaac Sim.
- (2) Create a *Cabinet* asset.
- (3) Randomize cabinet joint physics to vary friction and damping (simulating “stuck” or “heavy” conditions).
- (4) Expose joint torques and rigid-body force tensors to the Python interface as digitized sensor signals for the LLM.

Phase 3: DPTA Loop Implementation The experiment compares a standard LLM planner against the DPTA framework.

5.3.3.1. A. High-Level Planner The LLM receives the command “Open the cabinet” and decomposes it into sub-tasks:

$$\text{move_to(handle)} \rightarrow \text{grasp()} \rightarrow \text{pull}().$$

5.3.3.2. B. Penetrative Perception (Differentiator)

- **Baseline:** Standard planners attempt the DMP without torque feedback. High friction causes stalling or gripper slip, and the visual system may incorrectly retry the same action.
- **DPTA:**
 - (1) Feed digitized torque/force sequences into the LLM.
 - (2) Prompt includes expert knowledge: e.g., “High torque spikes with low velocity indicate a stuck mechanism.”
 - (3) LLM reasoning (inner monologue): Detects abnormal resistance and outputs an adaptation strategy (e.g., increase DMP gain, apply a wiggle motion).

Phase 4: Robust Learning with HER

- Run 4,096 parallel environments in Isaac Gym.
- Slippage or partial success is stored using Hindsight Experience Replay (HER). For example, failing to open the drawer but maintaining the grasp is relabeled as a success for *grasping*.
- The replay buffer refines the low-level controller, ensuring that the DMP can adapt to high-resistance situations.

5.4 Evaluation Metrics

- (1) **Success Rate:** Percentage of successful openings across friction values from 0.1 to 1.0. Hypothesis: Standard planners fail at high friction; DPTA maintains high success.
- (2) **Adaptation Time:** Steps or seconds until the LLM detects a “stuck” state and adapts.
- (3) **Token Efficiency:** Ability of raw torque/force sequences to inform the LLM about physical state (Stuck vs. Heavy vs. Locked).

5.5 Expected Outcome

Visual grounding indicates *what* to manipulate, whereas the penetrative perception layer determines *how* to manipulate dynamically. Human-Robot Collaboration provides the kinematic baseline (DMP) that code-generation alone cannot reliably produce for complex curves. The DPTA framework is expected to robustly adapt motion primitives to varying physical constraints that would otherwise cause standard planners to fail.

6. CONCLUSION

7. CONCLUSION

This work introduces the **Dynamic Penetrative Trajectory Adaptation (DPTA)** framework, a novel architecture designed to bridge the critical gap between high-level semantic planning and low-level dynamic control in embodied agents. While Large Language Models (LLMs) possess vast world knowledge capable of decomposing abstract tasks into mid-level plans, they inherently lack real-world experience and physical grounding, often resulting in plans that are plausible but not executable. **DPTA** addresses these limitations by synthesizing *Penetrative AI*, *Dynamic Movement Primitives (DMP)*, and *Hindsight Experience Replay (HER)* into a cohesive system.

First, DPTA advances the perception capabilities of robotic agents by adopting the *Penetrative AI* paradigm. Unlike traditional systems that rely primarily on visual-language models (VLMs) which may miss invisible physical properties, DPTA enables the LLM to interpret digitized sensor signals (such as force and torque data) directly. This allows the planner to deduce micro-level physical states—such as mechanism stiffness or surface friction—providing the deep physical context necessary for complex manipulation.

Second, we address the limitations of rigid code generation and open-loop planning. While approaches like *Code as Policies* can express logic and geometric reasoning, they struggle to synthesize complex, adaptive trajectory curves from scratch. By utilizing *Dynamic Movement Primitives (DMP)* acquired through *Human-Robot Collaboration (HRC)*, DPTA ensures that the robot’s motions are kinematically sound and human-like. The LLM acts as a high-level orchestrator that parameterizes these primitives based on the penetrative sensory feedback, rather than attempting to generate low-level control laws directly.

Finally, DPTA enhances system robustness through *Hindsight Experience Replay (HER)*. Traditional reinforcement learning struggles with the sparse, binary rewards typical of manipulation tasks. By allowing the agent to learn from failure—treating achieved states in failed trajectories as successful goals—HER significantly improves sample efficiency and policy robustness in high-dimensional state spaces.

In summary, DPTA moves beyond the “Say” and “Can” dichotomy by introducing a “Feel” and “Adapt” capability. By grounding semantic intent in penetrative sensory data and robust motion primitives, DPTA offers a path toward embodied agents that can not only plan logically but

also interact physically with the nuance and adaptability required for real-world deployment.

REFERENCES

- Andrychowicz, M., Wolski, F., Ray, A., Schneider, J., Fong, R., Welinder, P., McGrew, B., Tobin, J., Pieter Abbeel, O., and Zaremba, W. (2017). Hindsight experience replay. *Advances in neural information processing systems*, 30.
- Bhat, V., Kaypak, A.U., Krishnamurthy, P., Karri, R., and Khorrami, F. (2024). Grounding llms for robot task planning using closed-loop state feedback. *CoRR*.
- Brohan, A., Chebotar, Y., Finn, C., Hausman, K., Herzog, A., Ho, D., Ibarz, J., Irpan, A., Jang, E., Julian, R., et al. (2023). Do as i can, not as i say: Grounding language in robotic affordances. In *Conference on robot learning*, 287–318. PMLR.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., et al. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877–1901.
- Ebert, D.D., Van Daele, T., Nordgreen, T., Karekla, M., Compare, A., Zarbo, C., Brugnera, A., Øverland, S., Trebbi, G., Jensen, K.L., et al. (2018). Internet-and mobile-based psychological interventions: applications, efficacy, and potential for improving mental health. *European Psychologist*.
- Gal, Y. and Ghahramani, Z. (2016). Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In *international conference on machine learning*, 1050–1059. PMLR.
- Huang, W., Abbeel, P., Pathak, D., and Mordatch, I. (2022). Language models as zero-shot planners: Extracting actionable knowledge for embodied agents. In *International Conference on Machine Learning*.
- Huang, W., Xia, F., Xiao, T., Chan, H., Liang, J., Florence, P., Zeng, A., Tompson, J., Mordatch, I., Chebotar, Y., et al. (2023). Inner monologue: Embodied reasoning through planning with language models. In *Conference on Robot Learning*, 1769–1782. PMLR.
- Liang, J., Huang, W., Xia, F., Xu, P., Hausman, K., Ichter, B., Florence, P., and Zeng, A. (2023). Code as policies: Language model programs for embodied control. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE.
- Liu, H., Zhu, Y., Kato, K., Tsukahara, A., Kondo, I., Aoyama, T., and Hasegawa, Y. (2024). Enhancing the llm-based robot manipulation through human-robot collaboration. *IEEE Robotics and Automation Letters*, 9(8), 6904–6911.
- Strobel, V., Dorigo, M., and Fritz, M. (????). Llm2swarm: Robot swarms that responsively reason, plan, and collaborate through llms. In *NeurIPS 2024 Workshop on Open-World Agents*.
- Vemprala, S.H., Bonatti, R., Bucker, A., and Kapoor, A. (2024). Chatgpt for robotics: Design principles and model abilities. *Ieee Access*, 12, 55682–55696.
- Wang, Y.J., Zhang, B., Chen, J., and Sreenath, K. (2024). Prompt a robot to walk with large language models. In *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 1531–1538. IEEE.
- Xu, H., Han, L., Yang, Q., Li, M., and Srivastava, M. (2024). Penetrative ai: Making llms comprehend the physical world. In *Proceedings of the 25th International Workshop on Mobile Computing Systems and Applications*, 1–7.

Appendix A. A SUMMARY OF LATIN GRAMMAR

Appendix B. SOME LATIN VOCABULARY