Modern robot learning has increasingly embraced large, data-hungry models following recent successes in vision and language. While these approaches have shown impressive results, robots must physically interact with their surroundings under the unforgiving constraints of real-world physics and their embodiment. To make such interactions reliable, understanding *which* aspects of data actually matter for control, and embedding physics- and control-driven inductive biases into end-to-end learning frameworks are crucial. **My research goal is to develop physics-aware learning-and-control methods that let robots acquire scalable contact-rich skills across locomotion and manipulation and robustly perform them across tasks and environments.** In the long term, I hope to become a professor leading a research lab that builds on these foundations to develop principled, deployable embodied systems that extend human reach in hazardous and inaccessible environments.

Prior-guided Control and Learning: My first deep dive into research came during my junior year at IIT Madras, when I was selected for the IUSSTF Viterbi Program and joined Dr. Quan T. Nguyen's lab at USC. At the time, heavy reward shaping and task-specific tricks were the de facto paradigm for RL-based agile locomotion. We instead asked how priors from even imperfect dynamics models could systematically guide learning. My main contribution was designing a closed-loop state reference generator (oracle), based on a linear optimal control technique [1] and the standard centroidal dynamics model [2], that generates parametrised motion priors for a multi-modal RL policy. This simple linear model proved surprisingly powerful: the policy that was trained to track these references achieved parkour-like locomotion over structured terrain and generalised across task parameters. The resulting second-author paper [3], now available as a preprint, and its reviews taught me a great deal about writing, narrative clarity, and the importance of real-world evaluation.

Encouraged by the locomotion results, we were curious about the task generality of the framework. Initial experiments and ablations I led on extending the pipeline to uni-object loco-manipulation tasks paved way for designing oracles as hybrid automata with multi-modes. However, we observed serious local-optima–for instance, in a soccer task the robot would inadvertently kick the ball while "reaching," getting stuck in that mode. This motivated us to introduce a mode-transition reward to regularize switching, yielding cross-embodiment behaviors such as soccer-style dribbling and omnidirectional box manipulation, culminating in a co-authored IROS 2025 paper [4]. Together, these projects convinced me that structured priors can meaningfully shape learning while also highlighting a key limitation: hand-designed oracles do not scale gracefully to richer tasks. This has led us to currently explore automating oracle generation from human data and our attempts at transfer of policies to a custom hardware exposed many facets of sim-to-real gap, preventing zero shot transfer.

System-Identification to Physics-Aware Adaptation: Building on these observations of sim-to-real gaps, I joined Dr. Guanya Shi's LeCAR Lab at CMU with a clear focus: can we systematically identify the physical parameters that are critical and bring the simulator closer to the real system? Standard recipes like domain randomization often hurt training performance and, for contact-rich legged systems with hybrid dynamics, classical differentiable system-identification methods break down. In a project I led, we instead framed identification as a sampling-based optimization problem, using CMA-ES to recover key inertial and actuator parameters by minimizing trajectory prediction error. Early results on highly dynamic tasks exposed pronounced nonlinearities in the actuators, leading us to introduce a non-linear motor model inspired by [5] and jointly identify its parameters. This enabled zero-shot transfer without test-time PD tuning, with the same pipeline extending to humanoids. Another key factor in the identification process was the data: trajectories must sufficiently excite the identified parameters to allow accurate estimation. To address this, I designed an active exploration framework that optimizes the input commands of a multi-behavior policy to maximize Fisher information, yielding trajectories that are both hardware-safe and highly informative. This led to SPI-Active [6], a framework that achieved 42–63% performance improvements over baselines across several locomotion tasks and was accepted as an oral at CoRL 2025 (top 5.7%, **Highest Review Score**).

Despite SPI-Active's success in reducing robot modeling error, it remains limited to parametric uncertainties, and a policy trained in an updated simulator may still encode incorrect dynamics or observation models. This motivated a complementary line of work on policy-side adaptation, where I am currently exploring how to inject real-world information directly into the policy framework. Instead of relying on tedious and reward-sensitive real-world RL fine-tuning, I am co-training an auto-regressive forward dynamics model as an adaptation module, whose learned features condition the task policy, effectively grounding it through a physics-aware interface. I am also investigating various architectures and training objectives that enable the seamless transfer of policy performance by fine-tuning *only* the dynamics model using real-world rollouts.

A parallel but deeply related insight emerged when I contributed to the deployment of HDMI(ICRA 2026 submission) [7], a framework for learning humanoid loco-manipulation skills from monocular RGB videos. My insights on refining contact-based rewards and visualization tools were key to diagnosing failure modes and achieving successful sim-to-real transfer. However, extensive experimentation revealed a clear limitation: mismatched objects and environments accounted for the majority of failures, indicating a need to explicitly integrate them into future sim2real pipelines. Hence, *I envision building a unified framework that combines robot-dynamics adaptation with object/environment-side grounding to enable robust realization of loco-manipulation skills in unstructured settings.*

Interdisciplinary Research: Beyond legged robotics, I have pursued multi-agent coordinated control. At CMU, in Prof. Maxim Likhachev's group, I helped transform a conflict-based multi-agent planner into an end-to-end system for multi-quadruped coordination, including trajectory smoothing, low-level tracking, and custom-trained locomotion policies that remained stable under rapid command changes. This pipeline enabled synchronized motion of up to twelve heterogeneous quadrupeds in simulation, and I later built a modular sim-to-real framework to deploy it on four Unitree Go1 robots, resulting in a co-first-authored ICAPS 2025 demo [8] and an open-source codebase. This experience reinforced the need to tightly couple planning, control, and learning, and we are now extending the system toward capability-aware MAPF for interactive coordination across heterogeneous robot teams.

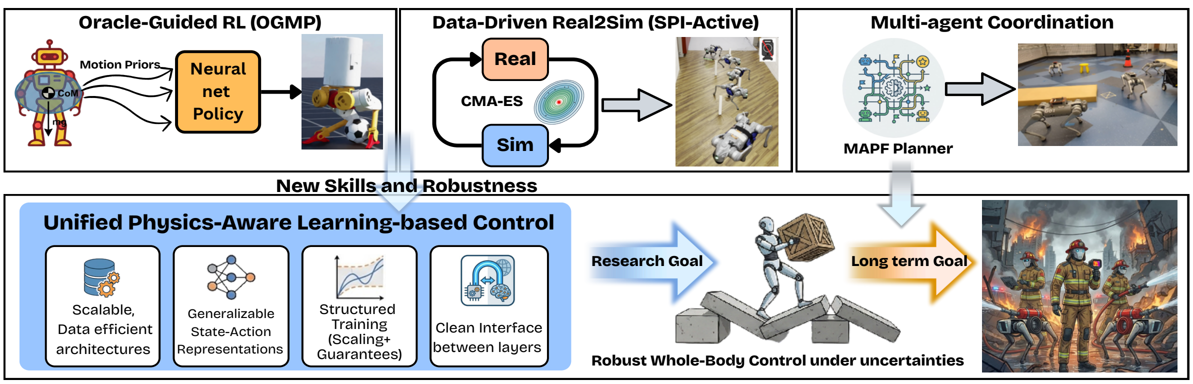

Figure 1: Prior research (Top) to Future research direction and Goals (Bottom)

Future Directions and Goals: My prior research has shown me structured recipes for bringing new capabilities and executing them effectively, while also preparing me to tackle fundamental questions in robot learning. Looking forward, I am broadly interested in the entire spectrum of algorithms and training paradigms at the intersection of learning and control that endow robots with contact-rich capabilities and then adapt those capabilities to new environments, tasks, and embodiments through a principled integration of data, physics, and control structure. I am particularly excited about whole-body control for legged robots, especially humanoids, which are increasingly emerging as a convergent form factor for physical automation. As depicted in Figure (1), concretely, I am interested in: **(i)** identifying model architectures that can exploit large-scale simulation while still being data-efficient with limited real-world experience; **(ii)** designing state–action representations that support generalization and adaptation; **(iii)** developing structured training pipelines that both admit control-theoretic guarantees and still benefit from scaling [9] and heterogeneous data sources; and **(iv)** creating clean interfaces between low-level controllers and large pretrained models trained on internet-scale data. These directions form the backbone of the research agenda I hope to pursue in graduate school.

References

[1] Masaki Murooka, Mitsuharu Morisawa, and Fumio Kanehiro. "Centroidal Trajectory Generation and Stabilization Based on Preview Control for Humanoid Multi-Contact Motion". In: *IEEE Robotics and Automation Letters* 7.3 (July 2022), pp. 8225–8232. ISSN: 2377-3774. DOI: 10.1109/lra.2022.3186515. URL: http://dx.doi.org/10.1109/LRA.2022.3186515.

[2] Jared Di Carlo, Patrick M. Wensing, Benjamin Katz, Gerardo Bledt, and Sangbae Kim. "Dynamic Locomotion in the MIT Cheetah 3 Through Convex Model-Predictive Control". In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2018, pp. 1–9. DOI: 10.1109/IROS.2018.8594448.

[3] Lokesh Krishna, **Nikhil Sobanbabu**, and Quan Nguyen. *OGMP: Oracle Guided Multi-mode Policies for Agile and Versatile Robot Control*. 2024. arXiv: 2403.04205 [cs.RO]. URL: https://arxiv.org/abs/2403.04205.

[4] Prashanth Ravichandar, Lokesh Krishna, **Nikhil Sobanbabu**, and Quan Nguyen. *Preferenced Oracle Guided Multi-mode Policies for Dynamic Bipedal Loco-Manipulation*. 2025. arXiv: 2410.01030 [cs.RO]. URL: https://arxiv.org/abs/2410.01030.

[5] Ruben Grandia, Espen Knoop, Michael A. Hopkins, Georg Wiedebach, Jared Bishop, Steven Pickles, David Müller, and Moritz Bächer. "Design and Control of a Bipedal Robotic Character". In: *Proceedings of Robotics: Science and Systems*. Delft, Netherlands, July 2024. DOI: 10.15607/RSS.2024.XX.103.

[6] **Nikhil Sobanbabu**, Guanqi He, Tairan He, Yuxiang Yang, and Guanya Shi. "Sampling-based System Identification with Active Exploration for Legged Sim2Real Learning". In: *9th Annual Conference on Robot Learning*. 2025. URL: https://openreview.net/forum?id=UTPBM4dEUS.

[7] Haoyang Weng, Yitang Li, **Nikhil Sobanbabu**, Zihan Wang, Zhengyi Luo, Tairan He, Deva Ramanan, and Guanya Shi. *HDMI: Learning Interactive Humanoid Whole-Body Control from Human Videos*. 2025. arXiv: 2509.16757 [cs.RO]. URL: https://arxiv.org/abs/2509.16757.

[8] Rishi Veerapaneni*, **Nikhil Sobanbabu***, Guanya Shi, Jiaoyang Li, and Maxim Likhachev. *Towards Unstructured MAPF: Multi-Quadruped MAPF Demo*. ICAPS 2025 Demonstrations Track. 2025. URL: https://icaps25.icaps-conference.org/program/demos-pdfs/ICAPS25-Demo_paper_8.pdf.

[9] Richard Sutton. "The Bitter Lesson". In: (2019). URL: http://www.incompleteideas.net/IncIdeas/BitterLesson.html.