Research Vision

Current methodologies in robot learning face two critical challenges that limit the capabilities of cyberphysical systems: expressing intentions as learning objectives (the intent-to-behavior gap) and maintaining controller behavior when moving from simulation to reality (the sim-to-real gap). These challenges are particularly acute in resource-constrained systems, where failures can lead to damage and where computational efficiency is paramount. My research introduces principled foundations for bridging these gaps through structured objective specification. By combining insights from programming languages, machine learning, and embedded systems, I enable practitioners to directly express their intentions while ensuring learned behaviors remains robust and efficient when deployed on real platforms—demonstrated through applications ranging from general-purpose robot control to state-of-the-art performance on racing quadrotors.

Research Impact and Current Work

Foundational Sim-to-Real Transfer: In RE+AL [1], my first research paper demonstrated the first RL-based controllers to consistently outperform well-tuned PID controllers on real racing quadrotors. Building on this, my recent work on anchor critics [2] allows for on-the-fly NN optimization, achieving near-50% power reduction while preventing catastrophic forgetting.

Reliable Real-World Deployment: In CAPS [3], a general power optimization technique, I demonstrated 80% power reduction while maintaining performance on real robots. This work has become a cornerstone reference in efficient reinforcement learning garnering over 80 citations.

Efficient Control for Resource-Constrained Robots: By breaking actor-critic symmetry [4], neural network controllers can be dramatically reduced in size (up to 99% reduction in number of parameters) with similar performance. Enabling deployment on resource-constrained platforms.

Future Research Program and Impact

My research agenda focuses on establishing Expressive Reinforcement Learning as a new subfield dedicated to faithfully translating high-level objectives into robust robot behaviors. I am developing Algebraic Policy Specification (APS) [5], a novel domain-specific language that generalizes traditional reward composition to enable direct expression of complex behavioral objectives. Initial results show a 600% improvement in sample efficiency compared to state-of-the-art methods. In the short term (2-5 years), I will create frameworks for composable reward specifications that generalize across different robotics domains while maintaining computational efficiency. Long-term, this work will establish theoretical foundations for a new generation of robot learning systems that can faithfully translate high-level intentions into reliable behaviors.

I am committed to open-source development and have released: a firmware for live adaptation on real robots [6], algorithms for Lyapunov-based stable control [7] and stochastic dynamics learning [8], and software for risk-aware planning [9]. Through these contributions and active mentorship of students in robotics research, including leading 11 students to second place in F1Tenth racing and mentoring graduate students to publication at ECCV 2024 [10], I aim to foster an environment where advances in robot learning translate directly to practical impact.

Bibliography

[1] B. Mabsout, S. Mysore, K. Saenko, and R. Mancuso, "How to train your quadrotor: A framework for consistently smooth and responsive flight control via reinforcement learning," *ACM Transactions on Cyber-Physical Systems*, vol. 5, no. 4, 2021, doi: 10.1145/3466618.

[2] B. Mabsout, S. Mysore, K. Saenko, and R. Mancuso, "Anchored Learning for On-the-Fly Adaptation," *arXiv preprint arXiv:2301.06987*, 2024, [Online]. Available: https://arxiv.org/abs/2301.06987

[3] B. Mabsout, S. Mysore, K. Saenko, and R. Mancuso, "Regularizing Action Policies for Smooth Control with Reinforcement Learning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2021. doi: 10.1109/ICRA48506.2021.9561138.

[4] S. Mysore, B. Mabsout, R. Mancuso, and K. Saenko, "Honey, I Shrunk The Actor: A Case Study on Preserving Performance with Smaller Actors in Actor-Critic Reinforcement Learning," in *IEEE Conference on Games (CoG)*, 2021. doi: 10.1109/CoG52621.2021.9619008.

[5] B. Mabsout, "Algebraic Policy Specification for Expressive Reinforcement Learning." [Online]. Available: https://github.com/bmabsout/AQS

[6] B. Mabsout, "SWANNFlight: Open-source firmware for reinforcement learning on quadrotors." [Online]. Available: https://github.com/bmabsout/AnchoredActorCritic

[7] B. Mabsout, "SystemDescent: Learning Lyapunov functions for stable control." [Online]. Available: https://github.com/bmabsout/SystemDescent

[8] B. Mabsout, "Swirls: Stochastic dynamics learning with GANs." [Online]. Available: https://github.com/bmabsout/swirls

[9] B. Mabsout, "Gaussian Racer: Risk-aware path planning using 2D-Gaussian mixtures." [Online]. Available: https://github.com/bmabsout/gaussian_racer_jax

[10] K. Sengupta, Z. Shagguan, S. Bharadwaj, S. Arora, E. Ohn-Bar, and R. Mancuso, "UniLCD: Unified Local-Cloud Decision-Making for Autonomous Mobile Robots," in *European Conference on Computer Vision (ECCV)*, 2024. [Online]. Available: https://unilcd.github.io/