Advancing Safe and Reactive Robot Skills in Dynamic Environments

Puze Liu
DFKI - Darmstadt, puze.liu@robot-learning.de

Dynamic tasks demand that robots respond and adapt rapidly to changing environments. Algorithms across all levels–perception, planning, control, optimization, and learning–must account for challenges such as environmental variations, computation constraints, time delays, and hardware constraints, often imposing higher system requirements to ensure robust performance. Remarkable success has been achieved in highly dynamic robotic tasks, such as Catching the Ball [1], Spherical Pendulum [2], Robot Juggling [3], and Table Tennis [4, 5]. In such tasks, robots are pushing to the limits of their hardware capabilities to achieve highly dynamic motions. **However, safety–one of the most critical aspects of robotics–is often overlooked or implicitly simplified in such scenarios.** For instance, robots are typically placed in obstacle-free environments, significantly simplifying the complexity of the task. On the other hand, when safety is the key research question, the robot's speed is often intentionally scaled down to values far below its theoretical limits [6, 7]. In fact, developing reactive motor skills and ensuring safe behaviors often present conflicting challenges. Slow, compliant motions can significantly reduce impact forces when interacting with obstacles, inherently enhancing safety. However, achieving reactive behavior requires high-performance tracking controllers and high-speed motion, which, in turn, increases the risk and severity of potential collisions. This fundamental trade-off underscores the difficulty of balancing agility and safety in dynamic robotic systems. Therefore, my research focuses on bridging the gap between dynamic motor skills while ensuring safety in robotic systems. The central question is: **How to develop reactive robotic motor skills to achieve high-speed tasks while ensuring safety in dynamic environments?**

Robot safety in dynamic tasks is highly challenging due to a wide range of environmental and task-specific factors: a) *Task Complexity*. The unpredictability of dynamic environments makes it difficult to cover all possible variations in algorithmic design. Defining both objectives and safety constraints is inherently complex and often leads to suboptimal solutions. b) *Sim-to-Real Gap*. The simplification of the simulated dynamic model [8] and the computational approximation [9] aggravate the discrepancy between simulation and reality, especially in high-speed tasks, making real-world deployment more challenging. c) *Reactiveness*. In dynamic tasks, robots must perform fast computations and reactive behavior. Obtaining safe motion within a short time frame is essential. d) *Physical Feasibility*. Learning algorithms should comply with physical constraints, including velocity continuity, actuator limits, and hardware limitations.

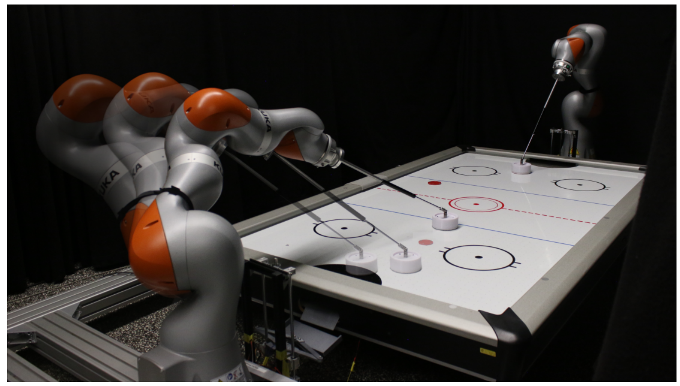

Fig. 1. Robot Air Hockey

I majorly use Robot Air Hockey (Figure 1) as a benchmark problem, as it requires the robot to perform high-speed motions in a restricted workspace, such as hitting and defending. Safety constraints are defined to avoid collisions with the table and respect joint position/velocity limits. Beyond robot air hockey, we further investigate the safety problem in the HRI (Human-Robot Interaction) and the navigation task, validating its effectiveness across diverse dynamic environments.

A. *Fast and Reactive Robot Skills via Optimization*

Safety problems are often defined as **constraints in the joint and task space**. Typical approaches, such as motion planning in task space, effectively address task space constraints. However, when Cartesian velocity is high, the resulting joint-space trajectory often becomes infeasible [10]. Instead, directly optimizing joint space trajectories while satisfying task space constraints is high-dimensional, nonlinear, and computationally expensive [11, 12].

To overcome this challenge in the robot air hockey task, we propose a **sequential optimization framework**. This framework decomposes the problem into four computationally efficient subproblems [13] that are low-dimensional and can be solved efficiently. In the first step, we search for an optimal hitting joint configuration by maximizing the measure of manipulability of the robot along the hitting direction. Then, we construct a linear programming problem to find the maximum hitting velocity considering joint velocity limits. Third, a trajectory planner is applied to obtain a task-space trajectory. Finally, we formulate a quadratic programming problem that exploits the robot's redundancy to find a physically feasible joint trajectory. This framework allows the robot to obtain a safe and feasible trajectory reactively (under 30 ms). While this framework generically focuses on the robot air hockey tasks, concepts, such as optimizing manipulability or exploiting redundancy, can be applied to other tasks.

While the decomposed sequential optimization framework improves computational efficiency, it suffers from the problem: the optimal solutions in the previous step may not be feasible in subsequent steps. To address this issue, in a collaborative effort, we developed a kinodynamic planner that eliminates the need for decomposition [14]. Our approach trains a neural network to directly infer the control points of a B-Spline. By constructing a differentiable loss function that encodes both task requirements (fast motions) and safety constraints, we obtain a neural planner that outputs high-speed motions while satisfying safety constraints. The inference time is less than 10ms, allowing us to achieve reactive behavior and dynamic hitting. We show in real-robot experiments that the trained kinodynamic planner results in lower tracking errors and a higher success rate than the baselines.

B. Safe Reinforcement Learning for Robotics

Previous approaches rely on prior knowledge of the system, task, and environment to design task-specific solutions. While effective and well-performed, these methods are task-dependent and lack generalizability. Instead, Reinforcement Learning (RL) offers a powerful framework for solving complex problems without domain-specific modeling. However, RL methods require accurate simulators and a huge amount of interactions, which makes the deployment of RL in real-world tasks challenging [15, 16, 17]. These challenges become even more pronounced in dynamic tasks, where capturing environmental variations along with real-world factors such as delays, observation noise, and disturbances, in simulation, is exceedingly difficult. Alternatively, developing safe and efficient RL algorithms that learn directly from real-world interactions will solve the issues effectively [18, 19, 20, 21].

In our initial work on this topic, we address the Safe Exploration (SE) problem, enabling robots to explore their environment while maintaining safety in the real world. Different from typical Safe RL algorithms that do not consider the constraint satisfaction during exploration [22, 23, 24], SE algorithms ensure safety at every step, such as constrained optimization [25, 26], Gaussian Processes [27], and backup policies [28]. Such methods are often algorithm-specific and lack generalizability across different RL frameworks.

To overcome these limitations, we propose a **safe exploration** method by constructing a constraint manifold. We then build an action space that allows agents to explore safely by determining the tangent space of the constraint manifold. This approach effectively transforms a constrained RL problem into an unconstrained one, making it compatible with any model-free RL algorithm while ensuring constraint satisfaction. We validate our approach in the Robot Air Hockey task and, to the best of our knowledge, demonstrate for the first time that a robotic system learns to solve a task directly in the real world from scratch while explicitly addressing safety constraints. Building upon this foundation, we show our approach can handle safety constraints in dynamic environments, such as collision avoidance in human-robot shared workspaces [29]. Notably, we have provided a grounded theoretical analysis and extensive validation of our approach across different robotic platforms–including manipulation, mobile robots, and quadrotors–demonstrating its robust generalizability [30].

C. Learning Constraint for HRI and Long-Term Safety

Lastly, I focus on constructing safety constraints in dynamic environments, the most critical challenge in safe motion planning and control [31, 32]. Our first attempt focuses on the collision avoidance problem by enforcing the distance between the robot and the obstacles bigger than a threshold. However, accurate distances are difficult to compute for objects with complex geometries or articulated structures. To address this, we propose a method that leverages neural networks with a robust inductive bias to train a Signed Distance Function (SDF). Our model offers precise distance computation for objects in proximity, meaningful level curves at the far end, and smooth and differentiable functions well-suited for reactive control or trajectory optimization. We validate our learned SDF model in building reactive whole-body controllers [33] and Safe RL policy in the bi-manipulation task and HRI task [29].

Another key challenge in constructing safety constraints is ensuring long-term safety, where robots must anticipate potential hazards and take early preventive actions—such as braking in advance to avoid collisions. Effectively tackling long-term safety requires predictive models capable of accounting for future outcomes, interactions, and environmental variations. To address these challenges, we proposed a data-driven approach that learns the long-term safety constraints, inspired by the value function in RL [34]. To further address the uncertainty raised by the system, we use the distributional RL to construct the safety constraint as a random variable, enabling the training of a risk-aware Safe RL algorithm.

FUTURE RESEARCH

In my future research, I will investigate the following topics:

a) Safe Hierarchical Policies: Solving complex robotic tasks usually requires a hierarchical policy structure. High-level policies focus on abstract concepts of the task, while low-level policies focus on the execution level[35, 36]. Robot safety should specified in different levels of the policy. However, high-level policies may generate infeasible commands for low-level policies to execute. I will to investigate how to build consistent hierarchical safe policies for robots.

b) Robot Safety in the Open World: Current existing safety methods are often designed for the closed environment, i.e., the task domain and environment variations are fixed in a certain domain. When deploying the robot in the open world, defining safe constraints to cover all environment scenarios is impractical. How to encode common knowledge (e.g., LLM, Knowledge Graph) in building safety specifications is one important problem to be explored in my future research.

References

[1] K. Dong, K. Pereida, F. Shkurti, and A. P. Schoellig, "Catch the ball: Accurate high-speed motions for mobile manipulators via inverse dynamics learning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 6718–6725.

[2] P. Klink, F. Wolf, K. Ploeger, J. Peters, and J. Pajarinen, "Tracking control for a spherical pendulum via curriculum reinforcement learning," *arXiv preprint arXiv:2309.14096*, 2023.

[3] K. Ploeger, M. Lutter, and J. Peters, "High acceleration reinforcement learning for real-world juggling with binary rewards," in *Conference on Robot Learning*. PMLR, 2021.

[4] D. Büchler, S. Guist, R. Calandra, V. Berenz, B. Schölkopf, and J. Peters, "Learning to play table tennis from scratch using muscular robots," *IEEE Transactions on Robotics*, vol. 38, no. 6, 2022.

[5] D. B. D'Ambrosio, S. Abeyruwan, L. Graesser, A. Iscen, H. B. Amor, A. Bewley, B. J. Reed, K. Reymann, L. Takayama, Y. Tassa *et al.*, "Achieving human level competitive robot table tennis," *arXiv preprint arXiv:2408.03906*, 2024.

[6] C. C. Beltran-Hernandez, D. Petit, I. G. Ramirez-Alpizar, T. Nishi, S. Kikuchi, T. Matsubara, and K. Harada, "Learning force control for contact-rich manipulation tasks with rigid position-controlled robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5709–5716, 2020.

[7] P. A. Lasota, T. Fong, J. A. Shah *et al.*, "A survey of methods for safe human-robot interaction," *Foundations and Trends® in Robotics*, vol. 5, no. 4, 2017.

[8] D. Baraff, "Physically based modeling: Rigid body simulation," *SIGGRAPH Course Notes, ACM SIGGRAPH*, vol. 2, no. 1, 2001.

[9] E. Todorov, "Convex and analytically-invertible dynamics with contacts and constraints: Theory and implementation in mujoco," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014.

[10] J. Z. Kolter and A. Y. Ng, "Task-space trajectories via cubic spline optimization," in *2009 IEEE international conference on robotics and automation*. IEEE, 2009.

[11] M. Stilman, "Task constrained motion planning in robot joint space," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2007.

[12] N. Ratliff, M. Zucker, J. A. Bagnell, and S. Srinivasa, "Chomp: Gradient optimization techniques for efficient motion planning," in *2009 IEEE international conference on robotics and automation*. IEEE, 2009.

[13] P. Liu, D. Tateo, H. Bou-Ammar, and J. Peters, "Efficient and reactive planning for high speed robot air hockey," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021.

[14] P. Kicki, P. Liu, D. Tateo, H. Bou-Ammar, K. Walas, P. Skrzypczyński, and J. Peters, "Fast kinodynamic planning on the constraint manifold with deep neural networks," *IEEE Transactions on Robotics*, 2023.

[15] J. Kober, J. A. Bagnell, and J. Peters, "Reinforcement learning in robotics: A survey," *The International Journal of Robotics Research*, vol. 32, no. 11, 2013.

[16] L. Brunke, M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig, "Safe learning in robotics: From learning-based control to safe reinforcement learning," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 5, 2022.

[17] J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, and S. Levine, "How to train your robot with deep reinforcement learning: lessons we have learned," *The International Journal of Robotics Research*, vol. 40, no. 4-5, 2021.

[18] L. Smith, J. C. Kew, X. B. Peng, S. Ha, J. Tan, and S. Levine, "Legged robots that keep on learning: Finetuning locomotion policies in the real world," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 1593–1599.

[19] L. Smith, I. Kostrikov, and S. Levine, "A walk in the park: Learning to walk in 20 minutes with model-free reinforcement learning," *arXiv preprint arXiv:2208.07860*, 2022.

[20] S. Ha, P. Xu, Z. Tan, S. Levine, and J. Tan, "Learning to walk in the real world with minimal human effort," *arXiv preprint arXiv:2002.08550*, 2020.

[21] J. Luo, Z. Hu, C. Xu, Y. L. Tan, J. Berg, A. Sharma, S. Schaal, C. Finn, A. Gupta, and S. Levine, "Serl: A software suite for sample-efficient robotic reinforcement learning," *arXiv preprint arXiv:2401.16013*, 2024.

[22] J. Achiam, D. Held, A. Tamar, and P. Abbeel, "Constrained policy optimization," in *International conference on machine learning*. PMLR, 2017.

[23] Y. Chow, O. Nachum, E. Duenez-Guzman, and M. Ghavamzadeh, "A lyapunov-based approach to safe reinforcement learning," *Advances in neural information processing systems*, vol. 31, 2018.

[24] C. Tessler, D. J. Mankowitz, and S. Mannor, "Reward constrained policy optimization," *arXiv preprint arXiv:1805.11074*, 2018.

[25] G. Dalal, K. Dvijotham, M. Vecerik, T. Hester, C. Paduraru, and Y. Tassa, "Safe exploration in continuous action spaces," *arXiv preprint arXiv:1801.08757*, 2018.

[26] R. Cheng, G. Orosz, R. M. Murray, and J. W. Burdick, "End-to-end safe reinforcement learning through barrier functions for safety-critical continuous control tasks," in *Proceedings of the AAAI conference on artificial intelligence*, vol. 33, 2019.

[27] F. Berkenkamp, M. Turchetta, A. Schoellig, and A. Krause, "Safe model-based reinforcement learning with stability guarantees," *Advances in neural information processing systems*, vol. 30, 2017.

[28] M. Alshiekh, R. Bloem, R. Ehlers, B. Könighofer, S. Niekum, and U. Topcu, "Safe reinforcement learning via shielding," in *Proceedings of the AAAI Conference

on Artificial Intelligence, vol. 32, 2018.

[29] P. Liu, K. Zhang, D. Tateo, S. Jauhri, Z. Hu, J. Peters, and G. Chalvatzaki, "Safe reinforcement learning of dynamic high-dimensional robotic tasks: navigation, manipulation, interaction," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2023.

[30] P. Liu, H. Bou-Ammar, J. Peters, and D. Tateo, "Safe reinforcement learning on the constraint manifold: Theory and applications," *IEEE Transactions on Robotics*, 2025.

[31] J. Ortiz, A. Clegg, J. Dong, E. Sucar, D. Novotny, M. Zollhoefer, and M. Mukadam, "isdf: Real-time neural signed distance fields for robot perception," *arXiv preprint arXiv:2204.02296*, 2022.

[32] D. Driess, J.-S. Ha, M. Toussaint, and R. Tedrake, "Learning models as functionals of signed-distance fields for manipulation planning," in *Proceedings of the 5th Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, vol. 164. PMLR, 08–11 Nov 2022. [Online]. Available: https://proceedings.mlr.press/v164/driess22a.html

[33] P. Liu, K. Zhang, D. Tateo, S. Jauhri, J. Peters, and G. Chalvatzaki, "Regularized deep signed distance fields for reactive motion generation," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2022.

[34] J. Günster, P. Liu, J. Peters, and D. Tateo, "Handling long-term safety and uncertainty in safe reinforcement learning," *Conference on Robot Learning*, 2024.

[35] Y. Song and D. Scaramuzza, "Learning high-level policies for model predictive control," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020.

[36] F. Stulp and S. Schaal, "Hierarchical reinforcement learning with movement primitives," in *2011 11th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2011.