Closing the Abstraction Gap: Adaptive Abstractions for Provably Safe Robot Planning

Disha Kamale, University of Michigan

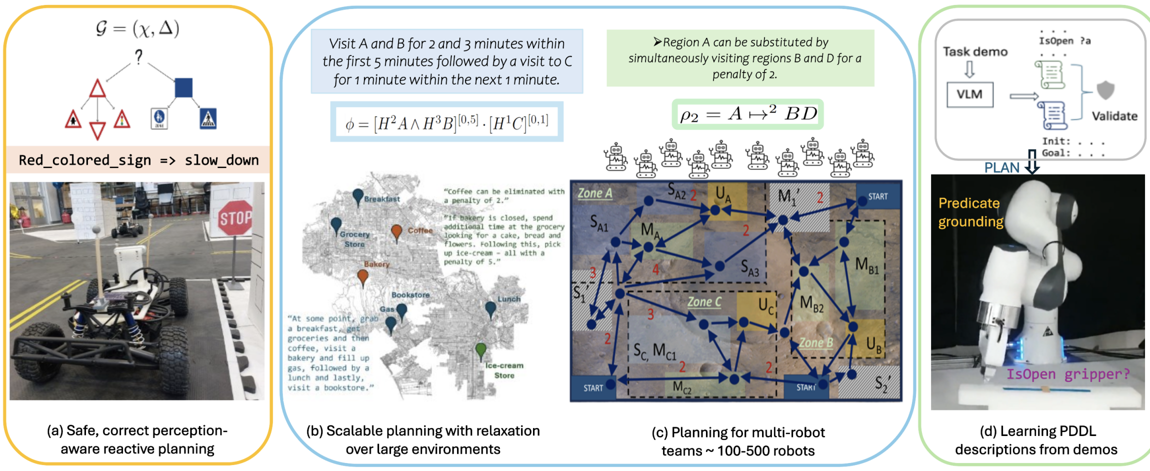

Fig. 1. a) Planning under incrementally improving perception: accounting partial symbolic information for effective decision-making [1–3], b)-c) formal control synthesis with relaxed satisfaction of complex goals over large environments and large robot teams [4–6], d) Learning task-level abstractions from demonstrations to enable long-horizon reasoning for complex manipulation tasks.

I. Introduction

Achieving reliable real-world autonomy requires robots to reason not only from raw sensory inputs and low-level control, but also over structured abstractions that connect perception, symbolic reasoning, and task specifications [7, 8]. These abstractions enable tractable long-horizon decision-making by capturing task-relevant semantics beyond raw observations. While learning-based approaches have demonstrated impressive progress from scene interpretation [9–11] to behavior cloning [12–15], these opaque approaches often struggle to generalize reliably or provide formal safety guarantees. Conversely, formal methods including temporal logic planning [16–18], reachability analysis, barrier certificates, verification, monitoring, and emerging neuro-symbolic approaches [19–21] offer mathematically grounded guarantees and have shown promise across domains from robotic manipulation to autonomous driving and healthcare. However, their effectiveness ultimately depends on the availability of suitable symbolic abstractions.

Existing research has explored this challenge from multiple directions. In perception-aware planning [22, 23], most approaches adopt a modular and sequential pipeline, typically treating symbolic knowledge as binary (either unavailable or fully known) rather than incrementally evolving as new observations arrive [24–28]. This limits their ability to utilize partial semantic information during execution, which is essential for real-time, safe decision-making. In formal control synthesis, automata-theoretic methods [29] provide expressive tools for specifying complex tasks, whereas sampling-based [16, 30, 31] and optimization-based [32, 17] control synthesis methods offer scalable solutions. Complementarily, work on learning task-level abstractions, such as Planning Domain Definition Language (PDDL) models [33] for automated planning and task-and-motion planning, has aimed to reduce manual modeling effort, but has largely relied on curated datasets and domain-specific structures. More recently, foundation models have opened new opportunities for generating symbolic abstractions using broad commonsense knowledge [34–36], yet integrating such learned representations with formal guarantees remains an open problem.

The aforementioned works implicitly rely on three key assumptions: (1) semantic knowledge of the environment is either complete or absent, rather than partial and evolving; (2) task specifications are fully feasible; and (3) symbolic task models are largely hand-crafted or fixed once constructed. Relaxing these assumptions is essential for enabling robots to operate reliably in dynamic, uncertain real-world settings.

My research addresses this gap through a unifying question: *How can robots construct and refine abstractions that link imperfect perception, learned task models, and complex, potentially infeasible goals while preserving provable guarantees?* I approach this through **abstraction-centric autonomy**: rather than relying on static, hand-engineered symbolic models, robots should continuously build and update structured representations from perception and experience. I pursue this vision along two complementary directions. First, I develop perception-aware formal planning methods that explicitly account for evolving symbolic knowledge, enabling correct-by-construction decision-making under partial information and possible infeasibility. Second, leveraging the recent advances in foundation models, I study how robots can automatically acquire symbolic planning models from demonstrations using visual-language information, reducing manual modeling effort while supporting scalable long-horizon reasoning. Together, these directions aim to close the loop between perception, abstraction, and formal decision-making.

This work sits at the intersection of formal methods, robot learning, and perception-aware planning. Whereas traditional approaches treat perception, model construction, and planning as loosely coupled stages with fixed symbolic representations, my research emphasizes adaptive abstractions that evolve with sensing, motion, and experience. By explicitly addressing partial knowledge, infeasible specifications, and learned symbolic structure, this work advances more robust, scalable, and interpretable planning frameworks, extending formal guarantees toward realistic deployment settings.

II. Research Contributions

Decision-making with imperfect symbolic perception At the core of this research thrust are the questions: 1) How to ensure that robots make correct decisions *in a timely manner*? 2) Under imperfect perception, how can any potentially relevant information be effectively utilized for safe planning? We model symbolic perception as progressively refining as the robot interacts with its environment, and introduce an incremental-resolution symbolic perception tree [1] - a principled abstraction that captures how environmental symbols and tracking accuracy improve over time. By integrating this perception hierarchy with motion control and assume–guarantee contracts, we formulate a Generalized Reactivity-1 (GR(1)) game that synthesizes provably correct reactive policies while explicitly accounting for partial symbolic information. This framework makes the evolution of the robot's knowledge explicit in the planning process, improving transparency and enabling formally grounded decision-making under realistic sensing conditions.

Building on this abstraction and moving beyond reactive planning, we extend the approach to resource-constrained exploration, where partial semantic knowledge actively guides information-gathering behavior [2]. We formulate this setting as an optimal-flow problem and develop a Mixed-Integer Linear Programming (MILP) method that synthesizes control policies online as the robot's understanding of the environment improves, enabling efficient task completion even without prior target information. Complementing these results, we further introduce a risk-aware planning framework [3] that quantifies the impact of perceptual uncertainty on decision-making, allowing robots to balance task progress against the likelihood and consequences of misperception. Together, these contributions establish a unified approach for incorporating evolving symbolic perception directly into formal planning and control, advancing both the reliability and practical deployability of autonomous robotic systems.

Optimal planning for infeasible missions Even under perfect perception, often, the complex missions assigned to robots may have some infeasible sub-requirements that may render control synthesis infeasible. To achieve meaningful satisfaction of tasks rather than simply giving up in such scenarios, it is essential to incorporate user preferences for task relaxations into the formal planning framework. Starting with a purely automata-based approach, our work in [4] incorporates user preferences into formal control synthesis framework and unifies several notions of relaxation existing in the literature e.g., minimum violation [37–40], deadline relaxation [41], etc. While this approach is highly expressive, it becomes computationally intractable for large robot teams and large environments. To alleviate this, we propose a novel, scalable flow-based MILP formulation to plan for homogeneous multi-robot teams [5]. Notably, our formulation is agnostic to team size. This approach brings together the strengths of expressivity and computational efficiency of automata and MILP-based synthesis, respectively, while avoiding their shortcomings of limited scalability and lack of explicit indication of progress. To further support planning in large environments, we introduce an A^*-based method for planning with relaxation that employs a novel specification-driven heuristic to guide the search. This substantially accelerates synthesis, enabling control plans for complex specifications to be generated within seconds for large city-scale graphs containing 10^6 edges, outperforming baselines by over 3x.

III. Ongoing and Future Directions

Ongoing: VLM-guided Abstraction generation. My current research focuses on autonomously generating formal task representations, such as PDDL-type descriptions, to enable long-horizon manipulation from demonstrations. Manually composing PDDL for complex tasks is error-prone and impractical for large domains. We leverage vision-language models to generate both PDDL domain and problem files, validated through VLM tool-use abilities. This approach produces structured PDDL representations that support feasible plan generation with classical planners, bridging the gap from raw demonstrations to actionable abstractions for principled decision-making.

Future Direction 1: Adaptive Perception Abstractions for Resource-constrained Planning. Real-world robots cannot afford to build complete semantic maps before acting. This direction develops representations that dynamically adjust their granularity based on mission objectives, task complexity, and resource constraints. Extending my prior work on incremental-resolution symbolic perception [1, 2] with foundation models' semantic understanding, these adaptive abstractions will enable informed, real-time decision-making in large, open-world environments without exhaustive pre-mapping.

Future Direction 2: Robust Long-horizon Planning via Skill Hierarchies and Failure Recovery. Long-horizon tasks require not only correct plans but resilience when those plans break down. This direction investigates learning reusable skill-level abstractions that structure complex tasks hierarchically, enabling efficient plan composition and reuse. Crucially, by integrating principled infeasibility handling - building on my work on specification relaxation - the resulting framework would support graceful failure recovery and online plan adaptation when environmental conditions or robot constraints change mid-execution.

REFERENCES

[1] Disha Kamale, Sofie Haesaert, and Cristian-Ioan Vasile. Cautious planning with incremental symbolic perception: Designing verified reactive driving maneuvers. *IEEE Intl. Conf. on Robotics and Automation*, 2023.

[2] Disha Kamale, Sofie Haesaert, and Cristian-Ioan Vasile. Energy-constrained active exploration under incremental-resolution symbolic perception. In *2023 62nd IEEE Conference on Decision and Control (CDC)*, pages 6863–6868. IEEE, 2023.

[3] Guangyi Liu, Vivek Pandey, Christoforos Somarakis, and Nader Motee. Risk of cascading failures in multi-agent rendezvous with communication time delay. In *2022 American Control Conference (ACC)*, pages 2172–2177. IEEE, 2022.

[4] Disha Kamale, Eleni Karyofylli, and Cristian-Ioan Vasile. Automata-based optimal planning with relaxed specifications. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 6525–6530. IEEE, 2021.

[5] Disha Kamale and Cristian-Ioan Vasile. Optimal control synthesis with relaxed global temporal logic specifications for homogeneous multi-robot teams. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 250–256. IEEE, 2024.

[6] Disha Kamale, Xi Yu, and Cristian-Ioan Vasile. A*-based temporal logic path planning with user preferences on relaxed task satisfaction. *arXiv preprint arXiv:2511.16844*, 2025.

[7] Cipriano Galindo, J-A Fernandez-Madrigal, and Javier Gonzalez. Improving efficiency in mobile robot task planning through world abstraction. *IEEE Transactions on Robotics*, 20(4):677–690, 2004.

[8] George Konidaris. On the necessity of abstraction. *Current opinion in behavioral sciences*, 29:1–7, 2019.

[9] Nicolas Schuler, Lea Dewald, Nick Baldig, and Jürgen Graf. From the laboratory to real-world application: evaluating zero-shot scene interpretation on edge devices for mobile robotics. In *International Conference on Innovative Techniques and Applications of Artificial Intelligence*, pages 301–315. Springer, 2025.

[10] Jianjun Ni, Yan Chen, Guangyi Tang, Jiamei Shi, Weidong Cao, and Pengfei Shi. Deep learning-based scene understanding for autonomous robots: A survey. *Intelligence & Robotics*, 3(3):374–401, 2023.

[11] Ruben Mascaro and Margarita Chli. Scene representations for robotic spatial perception. *Annual Review of Control, Robotics, and Autonomous Systems*, 8(1):351–377, 2025.

[12] Abhinav Kumar, Fan Yang, Sergio Aguilera Marinovic, Soshi Iba, Rana Soltani Zarrin, and Dmitry Berenson. Diffusing trajectory optimization problems for recovery during multi-finger manipulation. *arXiv preprint arXiv:2510.07030*, 2025.

[13] Yifan Zhang, Ruiping Wang, and Xilin Chen. Dynamic behavior cloning with temporal feature prediction: Enhancing robotic arm manipulation in moving object tasks. *IEEE Robotics and Automation Letters*, 2025.

[14] Shaopeng Liu, Chao Huang, and Hailong Huang. Behavior cloning-based active scene recognition via generated expert data with revision and prediction for domestic robots. *IEEE Transactions on Robotics*, 2025.

[15] Zengjie Zhang, Jayden Hong, Amir M Soufi Enayati, and Homayoun Najjaran. Using implicit behavior cloning and dynamic movement primitive to facilitate reinforcement learning for robot motion planning. *IEEE Transactions on Robotics*, 40:4733–4749, 2024.

[16] Cristian Ioan Vasile and Calin Belta. Sampling-based temporal logic path planning. In *2013 IEEE/RSJ international conference on intelligent robots and systems*, pages 4817–4822. IEEE, 2013.

[17] Gustavo A Cardona, Disha Kamale, and Cristian-Ioan Vasile. Mixed integer linear programming approach for control synthesis with weighted signal temporal logic. In *Proceedings of the 26th ACM International Conference on Hybrid Systems: Computation and Control*, pages 1–12, 2023.

[18] Gustavo A Cardona, Disha Kamale, and Cristian-Ioan Vasile. Stl and wstl control synthesis: A disjunction-centric mixed-integer linear programming approach. *Nonlinear Analysis: Hybrid Systems*, 56:101576, 2025.

[19] Chen Wang, Kaiyi Ji, Junyi Geng, Zhongqiang Ren, Taimeng Fu, Fan Yang, Yifan Guo, Haonan He, Xiangyu Chen, Zitong Zhan, et al. Imperative learning: A self-supervised neural-symbolic learning framework for robot autonomy. *arXiv preprint arXiv:2406.16087*, 2024.

[20] Xiaowu Sun and Yasser Shoukry. Neurosymbolic motion and task planning for linear temporal logic tasks. *IEEE Transactions on Robotics*, 2024.

[21] Xiao Li, Cristian-Ioan Vasile, and Calin Belta. Reinforcement learning with temporal logic rewards. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3834–3839. IEEE, 2017.

[22] Bunyo Okumura, Michael R James, Yusuke Kanzawa, Matthew Derry, Katsuhiro Sakai, Tomoki Nishi, and Danil Prokhorov. Challenges in perception and decision making for intelligent automotive vehicles: A case study. *IEEE Transactions on Intelligent Vehicles*, 1(1):20–32, 2016.

[23] Hongling Wang, Chengjin Zhang, Yong Song, and Bao Pang. A perception-driven exploration hierarchical simultaneous localization and mapping for mobile robots adapted to search and rescue environments. *Advances in Mechanical Engineering*, 10:1687814018765874, 2018.

[24] Georgios E. Fainekos, Hadas Kress-Gazit, and George J. Pappas. Temporal logic motion planning for mobile robots. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 2020–2025.

[25] Cameron Finucane, Gangyuan Jing, and Hadas Kress-Gazit. LTLMoP: Experimenting with language, temporal logic and robot control. In *2010 IEEE International Conference on Intelligent Robots and Systems*, pages

1988–1993.

[26] Hadas Kress-Gazit, Georgios E. Fainekos, and George J. Pappas. Where's waldo? sensor-based temporal logic motion planning. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 3116–3121.

[27] Yunus Emre Sahin, Petter Nilsson, and Necmiye Ozay. Multirobot coordination with counting temporal logics. *IEEE Transactions on Robotics*, 36(4):1189–1206, 2019.

[28] Yu Wang, Siddhartha Nalluri, and Miroslav Pajic. Hyperproperties for robotics: Planning via HyperLTL. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 8462–8468.

[29] Georgios E Fainekos. Revising temporal logic specifications for motion planning. In *2011 IEEE international conference on robotics and automation*, pages 40–45. IEEE, 2011.

[30] Yiannis Kantaros and Michael M Zavlanos. Sampling-based optimal control synthesis for multirobot systems under global temporal tasks. *IEEE Transactions on Automatic Control*, 64(5):1916–1931, 2018.

[31] Yiannis Kantaros and Michael M Zavlanos. Stylus*: A temporal logic optimal control synthesis algorithm for large-scale multi-robot systems. *The International Journal of Robotics Research*, 39(7):812–836, 2020.

[32] Kevin Leahy, Zachary Serlin, Cristian-Ioan Vasile, Andrew Schoer, Austin M Jones, Roberto Tron, and Calin Belta. Scalable and robust algorithms for task-based coordination from high-level specifications (scratches). *IEEE Transactions on Robotics*, 38(4):2516–2535, 2021.

[33] Constructions Aeronautiques, Adele Howe, Craig Knoblock, ISI Drew McDermott, Ashwin Ram, Manuela Veloso, Daniel Weld, David Wilkins Sri, Anthony Barrett, Dave Christianson, et al. Pddl—the planning domain definition language. 1998.

[34] Xiaohan Zhang, Zainab Altaweel, Yohei Hayamizu, Yan Ding, Saeid Amiri, Hao Yang, Andy Kaminski, Chad Esselink, and Shiqi Zhang. Dkprompt: Domain knowledge prompting vision-language models for open-world planning. *arXiv preprint arXiv:2406.17659*, 2024.

[35] Tom Silver, Ashay Athalye, Joshua B Tenenbaum, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. Learning neuro-symbolic skills for bilevel planning. *arXiv preprint arXiv:2206.10680*, 2022.

[36] Panagiotis Lymperopoulos and Liping Liu. Act-to-ground: A framework for symbol grounding in planning domains. In Leilani H. Gilpin, Eleonora Giunchiglia, Pascal Hitzler, and Emile van Krieken, editors, *Proceedings of The 19th International Conference on Neurosymbolic Learning and Reasoning*, volume 284 of *Proceedings of Machine Learning Research*, pages 777–795. PMLR, 08–10 Sep 2025. URL https://proceedings.mlr.press/v284/lymperopoulos25a.html.

[37] Jana Tumova, Luis I Reyes Castro, Sertac Karaman, Emilio Frazzoli, and Daniela Rus. Minimum-violation ltl planning with conflicting specifications. In *2013 American Control Conference*, pages 200–205. IEEE, 2013.

[38] Tichakorn Wongpiromsarn, Konstantin Slutsky, Emilio Frazzoli, and Ufuk Topcu. Minimum-violation planning for autonomous systems: Theoretical and practical considerations. In *2021 American Control Conference (ACC)*, pages 4866–4872. IEEE, 2021.

[39] Cristian-Ioan Vasile, Jana Tumova, Sertac Karaman, Calin Belta, and Daniela Rus. Minimum-violation scltl motion planning for mobility-on-demand. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1481–1488. IEEE, 2017.

[40] Lars Lindemann and Dimos V Dimarogonas. Control barrier functions for multi-agent systems under conflicting local signal temporal logic tasks. *IEEE control systems letters*, 3(3):757–762, 2019.

[41] Cristian-Ioan Vasile, Derya Aksaray, and Calin Belta. Time window temporal logic. *Theoretical Computer Science*, 691:27–54, 2017.