

Dynamic Closed-loop Replanning in Belief Space: Toward Handling Dynamically Changing Environments

Ali-akbar Agha-mohammadi, Saurav Agarwal, Suman Chakravorty, Nancy M. Amato

Abstract—Motion planning in belief space is a challenging problem due to the computational intractability of its exact solution. This problem becomes even more challenging in changing environments. This paper proposes a dynamic replanning scheme in belief space to handle changes in the environment (e.g., changes in the obstacle map), as well as unforeseen large deviations in the robot’s location (e.g., the kidnapped robot problem). The proposed method utilizes Feedback-based Information RoadMap (FIRM) framework as a substrate. However, it circumvents the need for belief stabilization in the FIRM framework and proposes a more efficient online planning scheme. We demonstrate the performance of the proposed method on a physical robot subject to large disturbances and environmental changes.

I. INTRODUCTION

This paper is concerned with the problem of Partially-Observable Markov Decision Process (POMDP) [4], which is a formal framework for sequential decision making under uncertainty. However, the POMDP problem is notorious for its computational intractability and developing approximate solutions for this problem is an ongoing research. Recently, the Feedback-based Information RoadMap (FIRM) framework [2] has provided a graph-based approach for belief space planning that significantly reduces the computational complexity of planning under uncertainty. In this paper we utilize the structure of FIRM to construct an efficient online replanning scheme.

Handling changes in the environment (e.g., obstacles), changes in the goal location, and large deviations in the robot’s location calls for online planning in uncertain, partially observable environments. One strategy to address this problem is an ability to dynamically replan in belief space. In this paper, we propose a principled rollout-based policy (ROP) algorithm based on the FIRM framework to construct an online stochastic replanning procedure.

Contributions: Contributions of this method over the original FIRM framework and receding horizon control-based (RHC-based) replanning schemes are as follows.

- An important contribution of the proposed work is its ability to bypass the belief stabilization process of the FIRM framework when there is no gain in stabilization. Thus, it generates plans with higher performance compared to the original FIRM framework.

Agha-mohammadi is with the Laboratory for Information and Decision Systems, MIT, Cambridge, MA, 02139. Agarwal and Chakravorty are with the Dept. of Aerospace Engineering, and Amato is with the Dept. of Computer Science and Engineering, Texas A&M University, TX, 77843, USA. Emails: aliagha@mit.edu {sauravag, schakrav, amato}@tamu.edu.



Fig. 1. A picture of robot (iRobot Create) in the operating environment. Landmarks can be seen on the walls.

- It provides a dynamic replanning scheme that can handle changes in the environment as well as large deviations in the system’s state.
- Compared to RHC-based methods in belief space, this method realizes a richer replanning scheme by providing a more accurate approximation in the planning horizon and incorporating a base cost-to-go beyond the horizon.
- An upper bound and a lower bound on the overall cost-to-go and success probability of the generated plan can be computed.
- Finally, we implement the proposed planner on a physical robotic system to demonstrate the performance and robustness of the method.

II. DYNAMIC REPLANNING IN BELIEF SPACE

FIRM graph: FIRM [2] is a “multi-query” graph in belief space. Each node of FIRM is a small region $B = \{b : \|b - \hat{b}\| \leq \epsilon\}$ around a sampled belief \hat{b} . We denote the i -th node by B^i and the set of nodes by $\mathbb{V} = \{B^i\}$. Each edge of FIRM is a local feedback controller, whose goal is to take the belief from its starting node into its end node. We denote the edge (controller) between nodes i and j by μ^{ij} and the set of edges by $\mathbb{M} = \{\mu^{ij}\}$. Associating appropriate costs and transition probabilities to the edges, we can solve a dynamic programming on the FIRM graph which leads to the optimal graph policy π^g that is a mapping from graph nodes to edges; i.e., $\pi^g : \mathbb{V} \rightarrow \mathbb{M}$. We denote the cost-to-go associated with π^g as J^g .

RHC in belief space: In the most common form of RHC [3] for stochastic systems, the system is approximated with a deterministic one by replacing the uncertain quantities with their typical values (e.g., maximum likelihood value). Then at every step the RHC scheme for deterministic systems solves an open-loop control problem (i.e., returns a sequence of actions $u_{0:T}$) over a fixed finite horizon T , executes only the first action u_0 , discards the remaining actions, and continues the same procedure in the next time step.

FIRM-based Rollout Policy (FIRM-ROP): Rollout Policy (ROP) approach [3] is a similar but more powerful replanning scheme than the described version of RHC in the following two senses. First, ROP does not approximate the system with a deterministic one and thus searches for a sequence of closed-loop policies (instead of open-loop controls) within the horizon. Second, ROP utilizes a suboptimal policy, called the “base policy,” to compute a cost-to-go function \tilde{J} that approximates the true cost-to-go beyond the horizon. In other words, at each step of the rollout policy scheme, the following closed-loop optimization is solved:

$$\pi_{0:T}(\cdot) = \arg \min_{\Pi_{0:T}} \mathbb{E} \left[\sum_{k=0}^T c(b_k, \pi_k(b_k)) + \tilde{J}(b_{T+1}) \right] \quad (1)$$

$$\begin{aligned} \text{s.t. } b_{k+1} &= \tau(b_k, \pi_k(b_k), z_k), \quad z_k \sim p(z_k|x_k) \\ x_{k+1} &= f(x_k, \pi_k(b_k), w_k), \quad w_k \sim p(w_k|x_k, \pi_k(b_k)) \end{aligned}$$

where $c(b, u)$ is the cost of taking action u at belief b . π_k is the policy at the k -th time step that belongs to Π_k . τ is the filtering equation that generates the next belief b_{k+1} based on the current belief b_k , control u_k , and observation z_{k+1} . x_k is the true system state and $p(z_k|x_k)$ is the likelihood distribution. f denotes the state evolution model, where w_k is the process noise. Note that in this formulation all future observations are taken into account.

Then, only the first control law π_0 is used to generate the control signal u_0 and the remaining policies are discarded. Similar to RHC, after applying the first control, a new sequence of policies is computed from the new point. In the FIRM-based rollout policy, we adopt the FIRM policy as the base policy of the rollout algorithm; i.e., $\tilde{J} = J^g$. Accordingly, the cost-to-go and success probability of the FIRM policy provides an upper and lower bounds for the cost and success probability of the FIRM-ROP. More details about this procedure can be found in [1].

Bypassing belief stabilization: In the original FIRM framework, at the end of each FIRM edge execution, the belief is stabilized to the end node of that edge. However, in the proposed framework due to the dynamic replanning procedure, at every step along the edge the method can decide to stop following the current edge and start going toward a new node. In other words if there is not enough gain in stabilization, the method will bypass it. To elaborate on this consider an example where we are only interested in minimizing the collision probability along the way to the goal. In that case, when the system is not in a narrow passage or not too close to obstacles, the method will bypass belief stabilization procedures as they do not affect the success probability of the mission. However, in narrow passages and close to obstacles, the method will lead to more conservative behaviors by stabilizing to FIRM nodes.

Changing environment and large deviations: In general, handling these cases in belief space is a big challenge as they require online updating of the planning structure in belief space. It is important to note that it is the unique graph structure of FIRM that makes such an update and replanning feasible in real-time. In case of obstacle map

changes, the graph structure of FIRM allows us to *locally* change collision probabilities without affecting the rest of the graph (i.e., properties of different edges on the graph are independent of each other). In case of large deviations, relying on the multi-query aspect of the FIRM graph, we can query the graph from the new deviated belief without re-evaluating graph edges. In belief planners that rely on forward search methods, collision probabilities and costs on *all* edges (number of possible edges is exponential in the number of underlying samples) need to be re-computed.

III. EXPERIMENTAL RESULTS

Next, we demonstrate the ability of the system to perform long-term tasks in a complex scenario that consists of visiting several goals (each time the robot reaches a goal, a user submits a new goal). The replanning ability allows the robot to change the plan online in belief space as the goal location changes. Moreover, the robot frequently encounters changes in the obstacle map (open/closed doors and new obstacles in the environment) as well as missing information sources and kidnapped robot situations. Thus, the robot frequently needs to perform a replanning operation in belief space to deal with such frequent changes. A 25-minute video of this run is recorded and available in [5] that shows the robot’s performance in this complex scenario. In this video, the robot faces three changes in the goal location, three changes in the door’s state (open/closed), several new obstacles in the environment, three kidnapping situations, and numerous failures of the sensory systems due to missing landmarks, blur in image, and etc.

IV. CONCLUSION

In this paper, we presented a dynamic replanning scheme in belief space. Such replanning is a key ability in handling discrepancies between real world models and computational models, changes in the environment and obstacles, and large deviations. We implemented this belief space planner on a physical system and demonstrated the robustness to such discrepancies that occur in practice. We believe this work provides an important step toward making POMDP methods applicable to real world robotic systems. Investigating the performance of the method on more challenging systems such as mobile manipulators is an interesting direction for future work.

REFERENCES

- [1] Aliakbar Agha-mohammadi, Saurav Agarwal, Aditya Mahadevan, Suman Chakravorty, Daniel Tomkins, Jory Denny, and Nancy Amato. Robust online belief space planning in changing environments: Application to physical mobile robots. In *IEEE Int. Conf. Robot. Autom. (ICRA)*, Hong Kong, China, 2014.
- [2] Aliakbar Agha-mohammadi, Suman Chakravorty, and Nancy Amato. FIRM: Sampling-based feedback motion planning under motion uncertainty and imperfect measurements. *International Journal of Robotics Research*, 33(2):268–304, Feb. 2014.
- [3] Dimitri Bertsekas. *Dynamic Programming and Optimal Control: 3rd Ed.* Athena Scientific, 2007.
- [4] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra. Planning and acting in partially observable stochastic domains. *Artificial Intelligence*, 101:99–134, 1998.
- [5] Video on Parasol Lab. webpage. <http://youtu.be/m3t3udm0ftu>.