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

Ali-akbar Agha-mohammadi[1], Saurav Agarwal[2], Suman Chakravorty[2], Nancy M. Amato[3]

[1]The Laboratory for Information and Decision Systems, MIT

[2]Dept. of Aerospace Engineering, Texas A&M University

[3]Dept. of Computer Science and Engineering, Texas A&M University

Introduction

Motion planning under motion and sensing uncertainty is an instance of the general problem of Partially-Observable Markov Decision Process (POMDP). The POMDP problem is a challenging problem due to the computational intractability of its exact solution. This problem becomes even more challenging in changing environments, because 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**.

Objective: Dynamic replanning in belief space

- A dynamic replanning scheme in belief space is proposed to handle
  - changes in the environment (e.g., obstacle map)
  - 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.
  - It generates plans with higher performance compared to the original FIRM framework since it is able to bypass the belief stabilization process
  - An upper bound and a lower bound on the overall cost-to-go and success probability of the generated plan is computed.
- Compared to RHC-based methods in belief space, this method
  - Considers all possible future observations
  - Incorporates a base cost-to-go beyond the horizon
- Finally, we implement the proposed planner on a physical robotic system to demonstrate the performance and robustness of the method.

Scenario

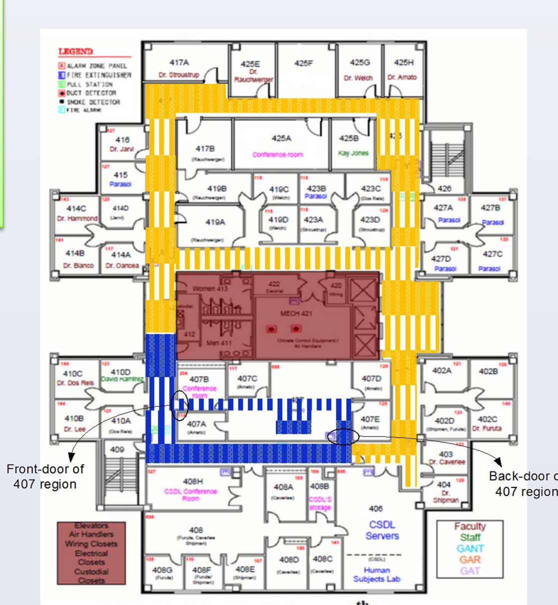
We conduct long-term missions consist of visiting several goals.

• A new task (goal) is submitted each time the robot reaches its current goal.

• The environment is changing (open/closed doors and moving people) as well as missing information

• Large deviations and Kidnapping situations.

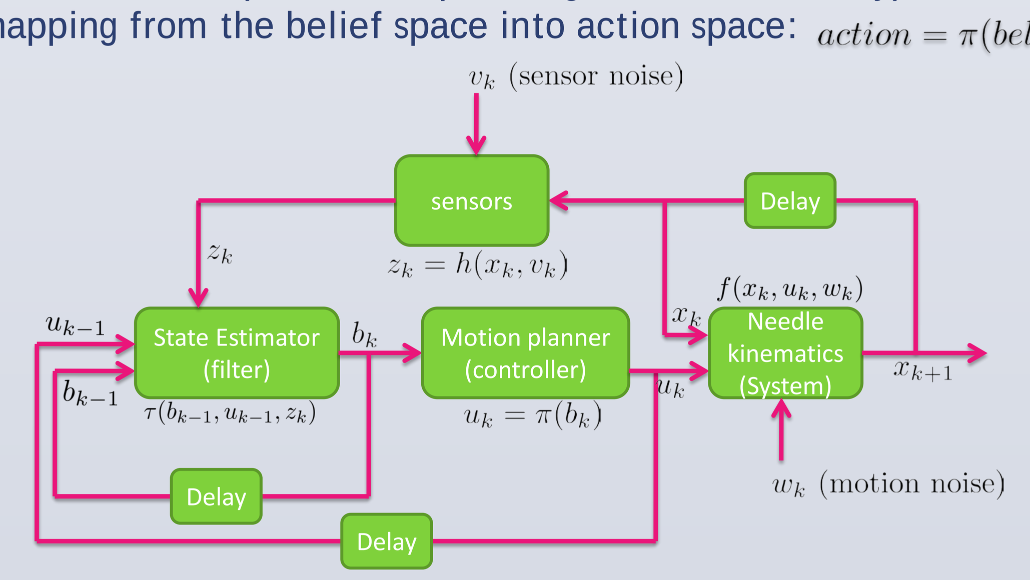
Thus, the robot needs to frequently perform online replanning in belief space to cope with these changes.



POMDP problem

• With noisy measurements, the only available information is the probability distribution over the state, which is called "belief" or "information-state": $b(x_k) = p(x_k | z_{0:k}, u_{0:k-1})$

• The feedback (solution of planning under uncertainty) is a mapping from the belief space into action space: $action = \pi(belief)$



$$J^{\pi}(b_0) = \mathbb{E}[\sum_{k=0}^{\infty} c(b_k, u_k)] = \mathbb{E}[\sum_{k=0}^{\infty} c(b_k, \pi(b_k))]$$

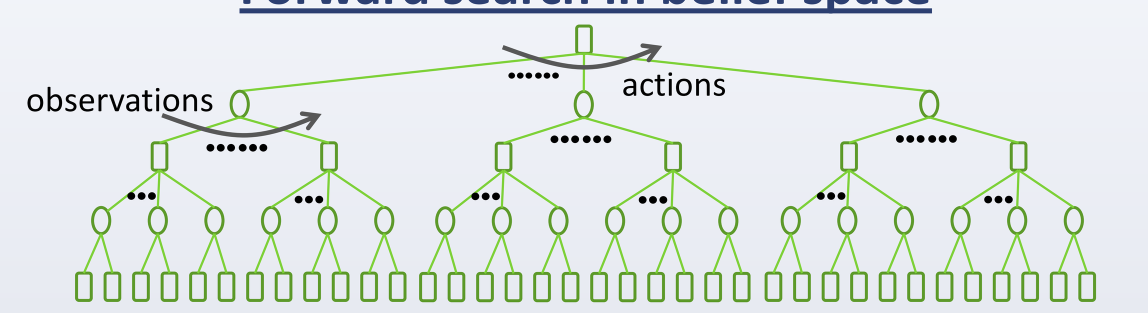
$$J(b_0) = \min_{\pi \in \Pi} J^{\pi}(b_0) \qquad \pi^* = \arg\min_{\pi \in \Pi} J^{\pi}(b_0)$$

$$J(b) = \min_u \{c(b,u) + \int_{\mathbb{B}} p(b'|b,u) J(b') db'\}, \quad \forall b \in \mathbb{B}$$

$$u^* = \pi(b) = \arg\min_u \{c(b,u) + \int_{\mathbb{B}} p(b'|b,u) J(b') db'\}$$

Cost in our experiments: $c(b_k, u_k) = \zeta_p \mathrm{tr}(P_k) + \zeta_u \|u_k\| + \zeta_T$

Forward search in belief space



Challenges in changing environment:

- Exponential growth
- Single-query

RHC in belief space

In the most common form of RHC the stochastic system is approximated with a deterministic system by replacing the uncertain quantities with their typical values (e.g., maximum likelihood value.)

$$u_{0:T} = \arg\min_{\mathbb{U}_{0:T}} \sum_{k=0}^{T} c(b_k^d, u_k)$$

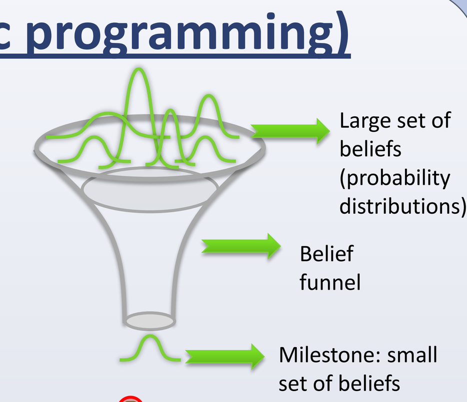
$$s.t. \quad b_{k+1}^d = \tau(b_k^d, u_k, z_{k+1}^{ml})$$

$$z_{k+1}^{ml} = \arg\max_z p(z | x_{k+1}^d)$$

$$x_{k+1}^d = f(x_k^d, u_k, 0)$$

Backward search (dynamic programming)

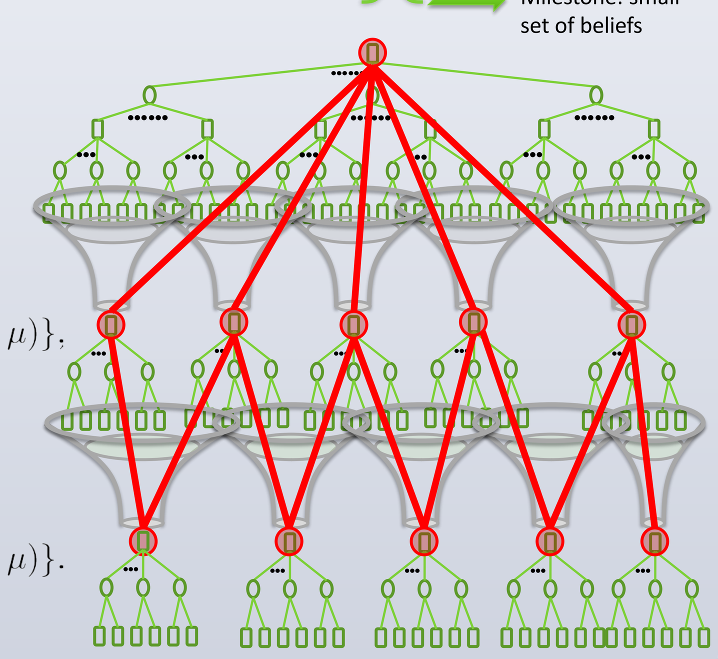
A belief funnel collapses a large set of beliefs to a small set of beliefs [1].



Dynamic programming based on funnels and milestones [1].

$$J^g(B^i) = \min_{\mu \in \mathbb{M}(i)} \{C^g(B^i, \mu) + J^g(F)\mathbb{P}^g(F|B^i, \mu) + \sum_j J^g(B^j)\mathbb{P}^g(B^j|B^i, \mu)\},$$

$$\pi^g(B^i) = \arg\min_{\mu \in \mathbb{M}(i)} \{C^g(B^i, \mu) + J^g(F)\mathbb{P}^g(F|B^i, \mu) + \sum_j J^g(B^j)\mathbb{P}^g(B^j|B^i, \mu)\}.$$



Rollout policy (ROP) in belief space

ROP is a dynamic replanning scheme where

• System is not approximated with a deterministic one within the horizon

• Cost-to-go beyond the horizon is approximated as the cost-to-go associated with a "base policy" [2]

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

$$s.t. \quad b_{k+1} = \tau(b_k, \pi_k(b_k), z_k), \quad z_k \sim p(z_k|x_k)$$

$$x_{k+1} = f(x_k, \pi_k(b_k), w_k), \quad w_k \sim p(w_k|x_k, \pi_k(b_k))$$

$$b_{\mathcal{T}+1} \in \cup_j B^j,$$

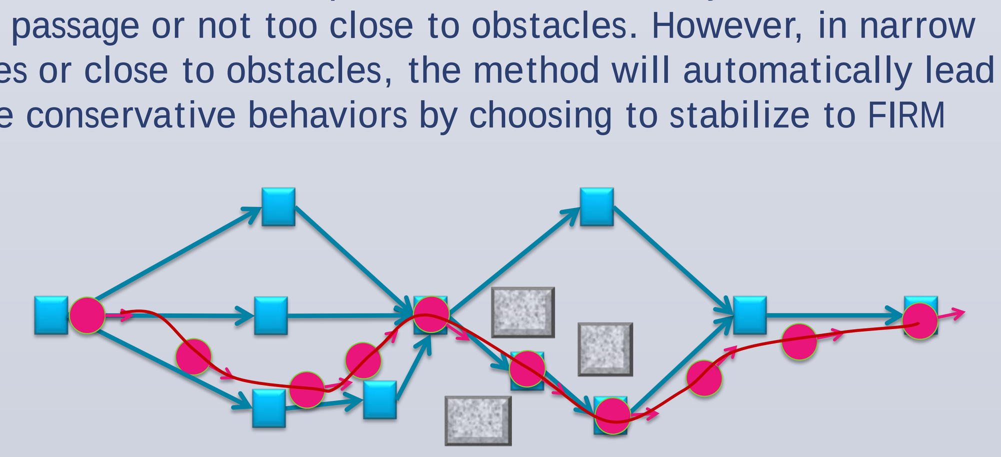
Feedback-based Information RoadMap (FIRM) is used as the "base policy" in belief space. Hence,

$$\widetilde{J}(b_{\mathcal{T}+1}) = J^g(B^j)$$

Bypassing stabilization

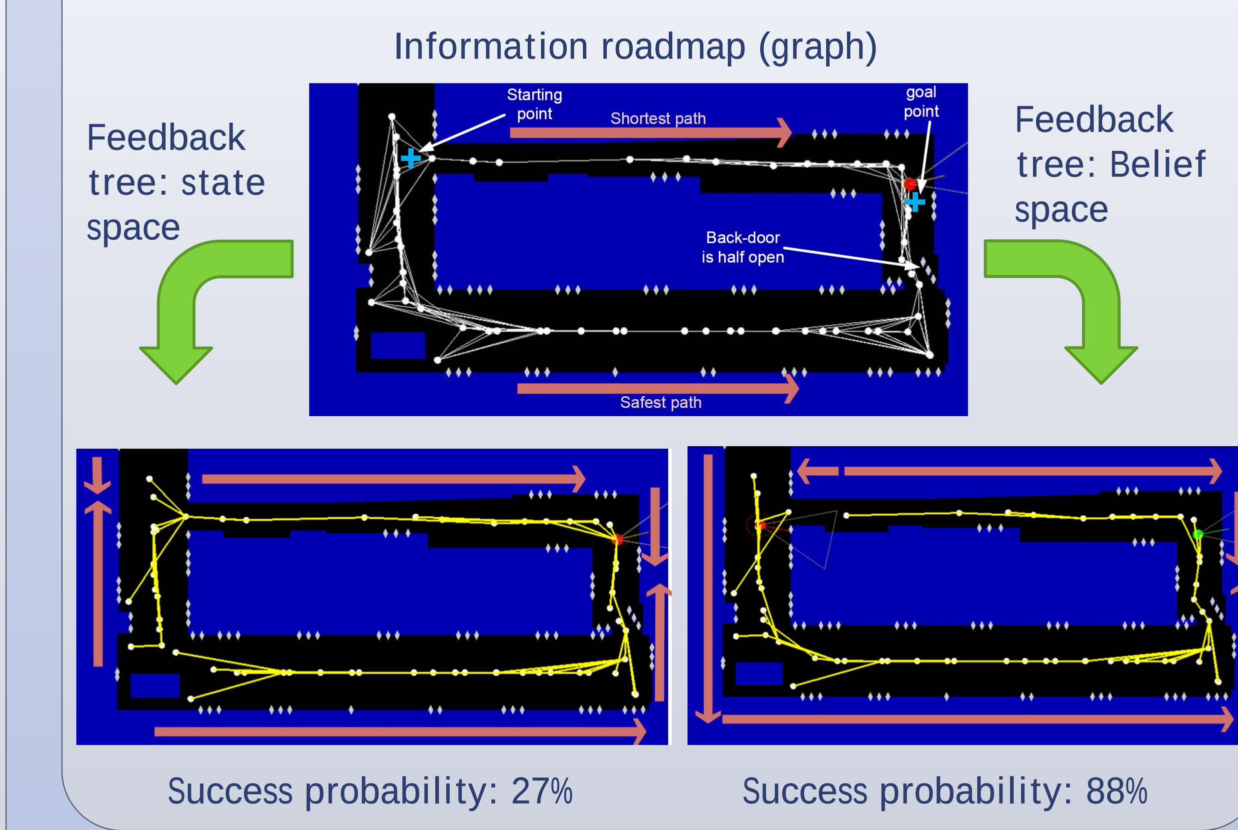
Due to the dynamic replanning procedure, at every step along the edge the method can decide to stop following the current edge and start going toward a new node. Thus, stabilization is bypassed if there is not enough gain in it.

For example, where we are only interested in minimizing the collision probability along the way to the goal, the method will bypass belief stabilization procedures when the system is not in a narrow passage or not too close to obstacles. However, in narrow passages or close to obstacles, the method will automatically lead to more conservative behaviors by choosing to stabilize to FIRM nodes.



Base policy: Feedback Tree

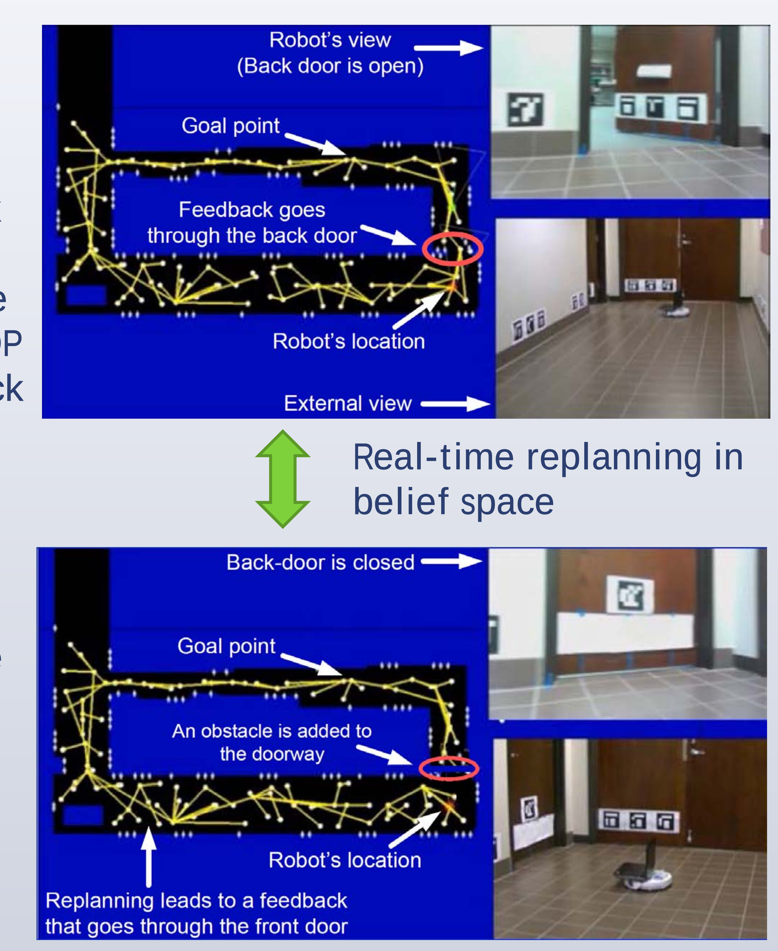
Base policy resulting from a FIRM is a feedback tree, which is a spanning tree of the underlying graph rooted in the goal node.



Replanning in Changing Environment

At every node the robot re-evaluates only the next (or the next few) edge(s) on the feedback tree. If there is a significant change in the collision probabilities, DP is re-solved and feedback tree is computed.

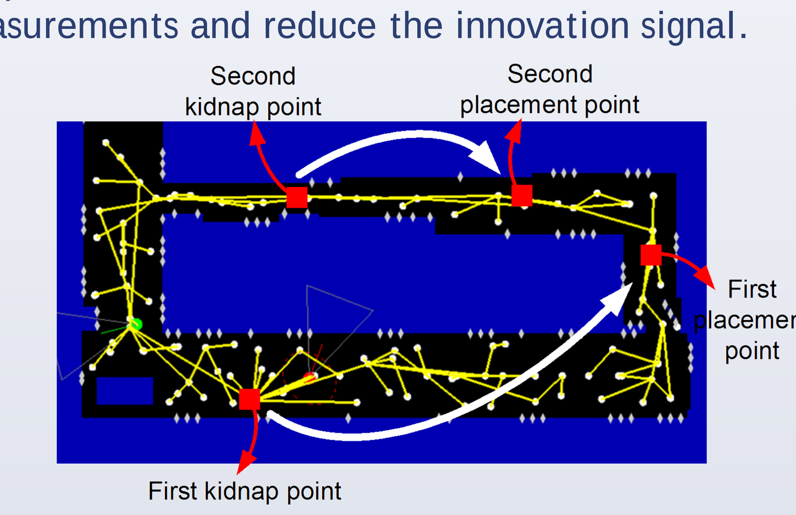
The graph structure allows us to *locally* update edges and hence perform real-time replanning.



Replanning after kidnapping

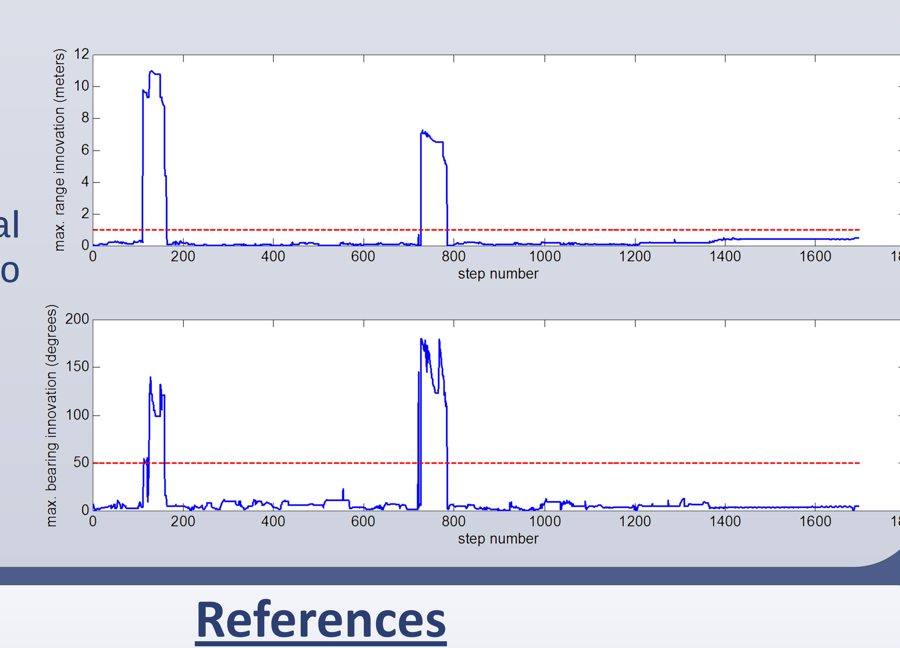
Kidnapping (large deviation) is detected by monitoring the innovation signal. In case of large deviations, we enter the Information Gathering Mode (IGM), where the robot takes conservative actions to get some known measurements and reduce the innovation signal.

Example: Two kidnappings on the way to goal



After reducing the innovation signal, we perform real-time replanning in belief space from the new deviated belief. Again, due to the graph structure that is spread in the belief space, the only required computation is to evaluate the cost of edges that connect the new starting belief to the neighboring FIRM nodes.

Innovation signal corresponding to the above example



References

[1] Ali-akbar Agha-mohammadi, Suman Chakravorty, Nancy M. Amato, "**FIRM: Sampling-based feedback motion planning under motion uncertainty and imperfect measurements**". International Journal of Robotics Research (2014)

[2] Ali-akbar Agha-mohammadi, Saurav Agarwal, Aditya Mahadevan, Suman Chakravorty, Daniel Tomkins, Jory Denny, Nancy M. Amato, "**Robust Online Belief Space Planning in Changing Environments: Application to Physical Mobile Robots,**" In Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Hong Kong, China, May 2014.

[3] Video of dynamic replanning in belief space on a physical robot in a changing environment. https://www.youtube.com/watch?v=6cKzcfVDes8