Motion Planning via Bayesian Learning in the Dark*

Authors: Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki

Presenters: Zahra Habibzadeh, Yara Mohammadi

Accepted for spotlight presentation at the Workshop on Machine Learning for Motion Planning ICRA 2021

Outline

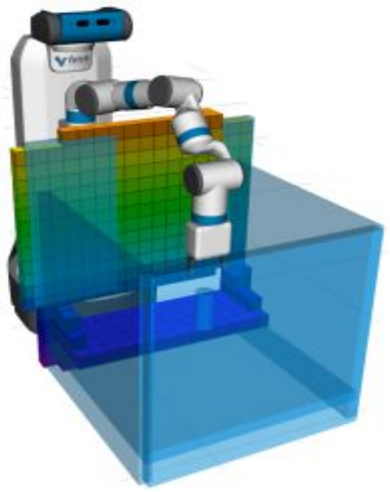
- Problem Definition
- Expression of necessity
- Previous Works
- Method
- Evaluation
- Conclusion
- References

Problem Definition

- In many applications ranging from robotic manipulation to autonomous driving, **motion planning** is a core problem

Problem Definition

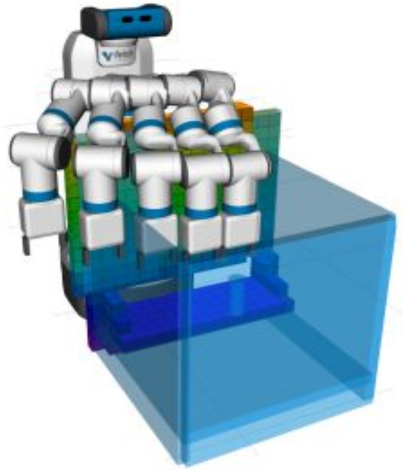
- It is tasked with picking an object that is located inside a box



[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Expression of necessity

- Complete knowledge of the robot's environment is required
- Limitations of robots’ sensors
- Human Guidance



[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Previous Works

- Robust motion planning using markov decision processes and quadtree decomposition **(MDP)** [1]
- Motion planning under uncertainty using iterative local optimization in belief space **(POMDPs)** [2]

Table 1. Overview of MDP and POMDP

<table>
<tr><th colspan="2" rowspan="2">Overview</th><th>Do we have control over the state transitions?</th></tr>
<tr><th>YES</th></tr>
<tr><td rowspan="2">Are the states completely observable?</td><td>YES</td><td>MDP
Markov Decision Process</td></tr>
<tr><td>NO</td><td>POMDP
Partially Observable Markov Decision Process</td></tr>
</table>

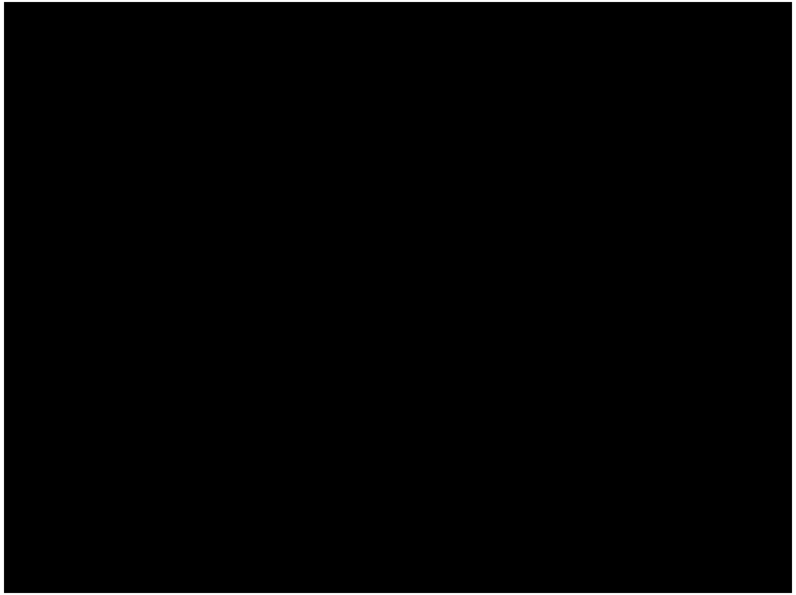
Previous Works

- LQG-MP: Optimized Path Planning for Robots with Motion Uncertainty and Imperfect State Information [3]
  - obstacle locations and shapes

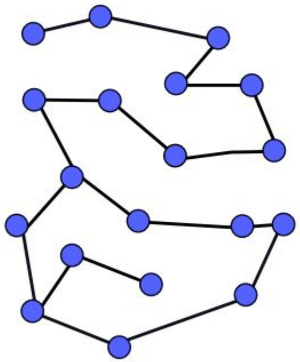
- Motion Planning for Manipulators in Unknown Environments with Contact Sensing Uncertainty [4]
  - limitations in sensing
  - The robot is allowed to make contact with the environment.

Previous Works

- Learning reward functions by integrating human demonstrations and preferences [5]
  - discrete state-spaces



Method



Roadmap G

Algorithm 1: BLIND

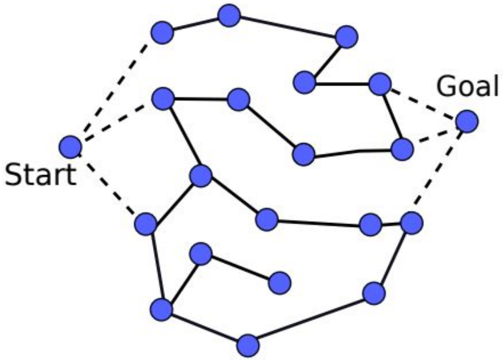
input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

```
G' ← ConnectToRoadmap (G, start, goal) ;
for i = 1,...,max_q do
    P ← GuidanceSearch (G') ;
    T ← GuidedMP (W, P, start, goal) ;
    if HumanAccepts (T) then
        Return T;
    else
        (D+, D−) ← Critique (T) ;
        Pr(R|D+, D−) ← BIRL (G', (D+, D−)) ;
        G' ← UpdateRoadmap (Pr(R|D+, D−)) ;
Return ∅ ;
```

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method



Roadmap G

Algorithm 1: BLIND

---

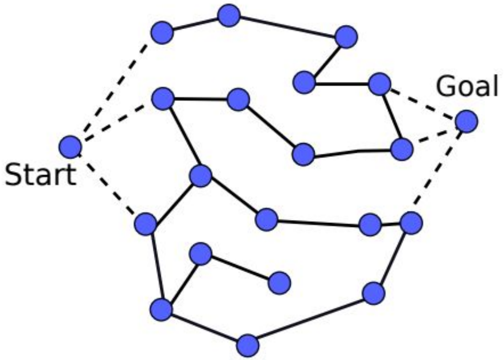
input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

```
G' ← ConnectToRoadmap (G, start, goal) ;
for i = 1,...,max_q do
    P ← GuidanceSearch (G') ;
    T ← GuidedMP (W, P, start, goal) ;
    if HumanAccepts (T) then
        Return T;
    else
        (D+, D−) ← Critique (T) ;
        Pr(R|D+, D−) ← BIRL (G', (D+, D−)) ;
        G' ← UpdateRoadmap (Pr(R|D+, D−)) ;
Return ∅ ;
```

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method



Roadmap G

Algorithm 1: BLIND

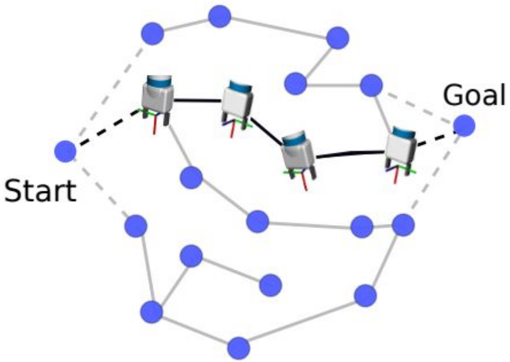
input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

$\mathcal{G}$' $\leftarrow$ `ConnectToRoadmap` ($\mathcal{G}$, start, goal) ;
for i = $1,\ldots,max_q$ **do**
  $P \leftarrow$ `GuidanceSearch` ($\mathcal{G}$') ;
  $\mathcal{T} \leftarrow$ `GuidedMP` ($\mathcal{W}$, P, start, goal) ;
  if `HumanAccepts` $(\mathcal{T})$ **then**
    Return $\mathcal{T}$;
  else
    $(D^+, D^-) \leftarrow$ `Critique` ($\mathcal{T}$) ;
    $\Pr(R|D^+, D^-) \leftarrow$ `BIRL` ($\mathcal{G}$', (D^+, D^-)) ;
    $\mathcal{G}$' $\leftarrow$ `UpdateRoadmap` ($\Pr(R|D^+, D^-)$) ;
Return $\emptyset$;

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method



Roadmap G

Algorithm 1: BLIND

input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

$\mathcal{G}$' $\leftarrow$ `ConnectToRoadmap` ($\mathcal{G}$, start, goal) ;

for i = $1,\ldots,max_q$ **do**

  $P \leftarrow$ `GuidanceSearch` ($\mathcal{G}$') ;

  $\mathcal{T} \leftarrow$ `GuidedMP` ($\mathcal{W}$, P, start, goal) ;

  if `HumanAccepts` ($\mathcal{T}$) **then**

    Return $\mathcal{T}$;

  else

    $(D^+, D^-) \leftarrow$ `Critique` ($\mathcal{T}$) ;

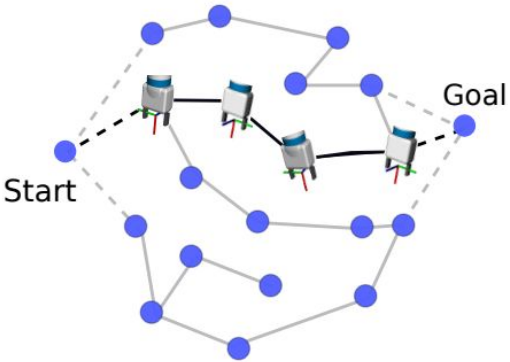
    $\Pr(R|D^+, D^-) \leftarrow$ `BIRL` ($\mathcal{G}$', (D^+, D^-)) ;

    $\mathcal{G}$' $\leftarrow$ `UpdateRoadmap` ($\Pr(R|D^+, D^-)$) ;

Return $\emptyset$;

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method



Roadmap G

Algorithm 1: BLIND

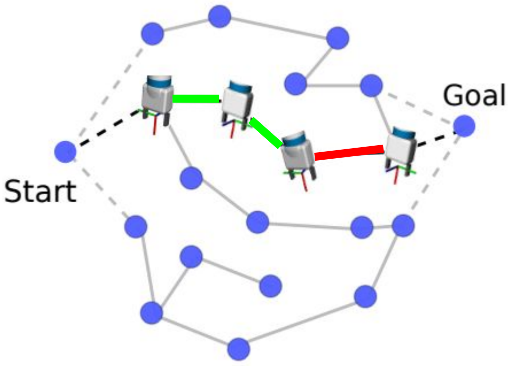
input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

$\mathcal{G}' \leftarrow$ `ConnectToRoadmap` ($\mathcal{G}$, start, goal) ;
for i = $1, \ldots, max_q$ **do**
  $P \leftarrow$ `GuidanceSearch` ($\mathcal{G}'$) ;
  $\mathcal{T} \leftarrow$ `GuidedMP` ($\mathcal{W}$, P, start, goal) ;
  if `HumanAccepts` ($\mathcal{T}$) **then**
    Return $\mathcal{T}$;
  else
    $(D^+, D^-) \leftarrow$ `Critique` ($\mathcal{T}$) ;
    $\Pr(R|D^+, D^-) \leftarrow$ `BIRL` ($\mathcal{G}'$, (D^+, D^-)) ;
    $\mathcal{G}' \leftarrow$ `UpdateRoadmap` ($\Pr(R|D^+, D^-)$) ;
Return $\emptyset$;

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method



Roadmap G

Algorithm 1: BLIND

input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

$\mathcal{G}$' ← `ConnectToRoadmap` ($\mathcal{G}$, start, goal) ;
for i = $1,\ldots,max_q$ **do**
  P ← `GuidanceSearch` ($\mathcal{G}$') ;
  $\mathcal{T}$ ← `GuidedMP` ($\mathcal{W}$, P, start, goal) ;
  if `HumanAccepts` ($\mathcal{T}$) **then**
    Return $\mathcal{T}$;
  else
    (D^+, D^-) ← `Critique` ($\mathcal{T}$) ;
    $\Pr(R|D^+, D^-)$ ← `BIRL` ($\mathcal{G}$', (D^+, D^-)) ;
    $\mathcal{G}$' ← `UpdateRoadmap` ($\Pr(R|D^+, D^-)$) ;
Return $\emptyset$;

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method: Key novelties

Algorithm 1: BLIND

input : Roadmap $\mathcal{G}$, incomplete workspace $\mathcal{W}$, start, goal, max_q

output : Collision-free Trajectory $\mathcal{T}$ or $\emptyset$ when fail

$\mathcal{G}$' $\leftarrow$ `ConnectToRoadmap` ($\mathcal{G}$, start, goal) ;

for i = $1,\ldots,max_q$ **do**

  $P \leftarrow$ `GuidanceSearch` ($\mathcal{G}$') ;

  $\mathcal{T} \leftarrow$`GuidedMP` ($\mathcal{W}$, P, start, goal) ;

  if `HumanAccepts` ($\mathcal{T}$) **then**

    Return $\mathcal{T}$;

  else

    $(D^+, D^-) \leftarrow$`Critique` ($\mathcal{T}$) ;

    $\Pr(R|D^+, D^-) \leftarrow$`BIRL` ($\mathcal{G}$', (D^+, D^-)) ;

    $\mathcal{G}$' $\leftarrow$`UpdateRoadmap` ($\Pr(R|D^+, D^-)$) ;

Return $\emptyset$;

Using the Roadmap to guide motion planning

Learning safety from human interactions

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Method: Using the task model to guide motion planning

- Optimization based motion planner (Guided TrajOpt)

$$
\begin{aligned}
\underset{x_0,\ldots,x_T}{\text{minimize}} \quad & \sum_{t=0}^{T-1} \|x_{t+1} - x_t\|^2 && \text{(1a)}\\
\text{subject to} \quad & x_0 = x_{St}, && \text{(1b)}\\
& x_T = x_G, && \text{(1c)}\\
& \mathrm{sd}(A_{it}, O_j) \geq d_s \ \forall i, j, t && \text{(1d)}\\
& F_k^{-1}\mathrm{FK}(x_\tau) = 0 \ \forall (k, \tau) \in P && \text{(1e)}
\end{aligned}
$$

3 $P \leftarrow$ `GuidanceSearch` $(\mathcal{G}')$;

4 $\mathcal{T} \leftarrow$`GuidedMP` $(\mathcal{W}, P, \text{start}, \text{goal})$;

Method: Learning safety from human interaction

$$R(v_1, v_2) = w \cdot \phi(v_1, v_2)$$

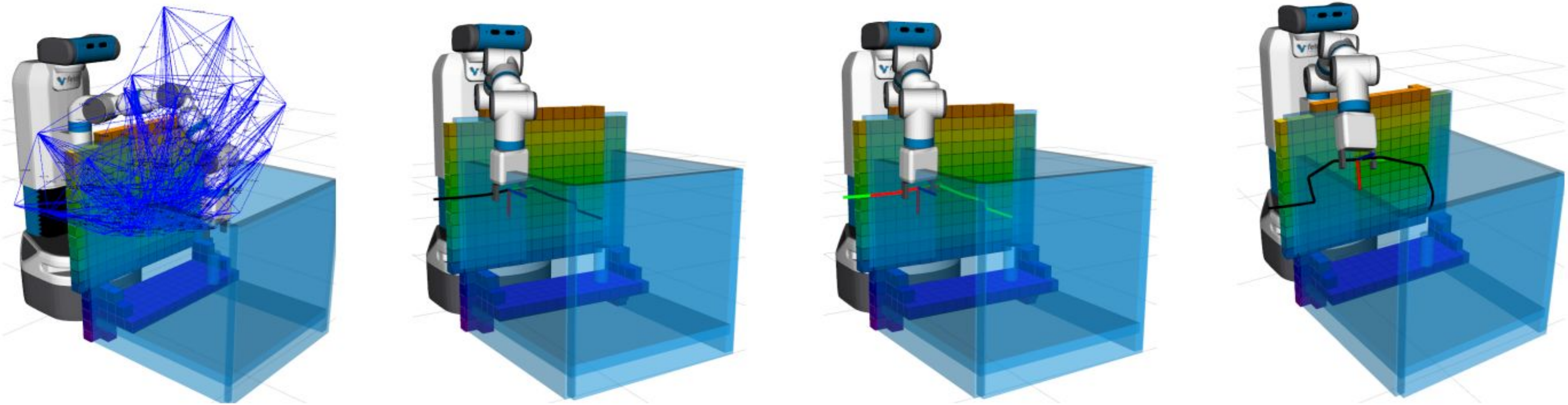
$$\begin{aligned}
\phi(v_a, v_b) &= [\phi_1, \phi_2, \phi_3, \phi_4, \phi_5] \\
\phi_1(v_a, v_b) &= |x_a + x_b| \\
\phi_2(v_a, v_b) &= |y_a + y_b| \\
\phi_3(v_a, v_b) &= |z_a + z_b| \\
\phi_4(v_a, v_b) &= \mathbb{1}_{(v_a, v_b) \in M} \\
\phi_5(v_a, v_b) &= \mathbb{1}_{(v_a, v_b) \notin M}
\end{aligned}$$

$(D^+, D^-) \leftarrow$ `Critique` $(\mathcal{T})$;

$\mathsf{Pr}(R|D^+, D^-) \leftarrow$ `BIRL` $(\mathcal{G}', (D^+, D^-))$;

$\mathcal{G}' \leftarrow$ `UpdateRoadmap` $(\mathsf{Pr}(R|D^+, D^-))$;

Method



a) Keypose Roadmap   b) Unsafe Guidance   c) Human Critique   d) Safe Guidance

[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

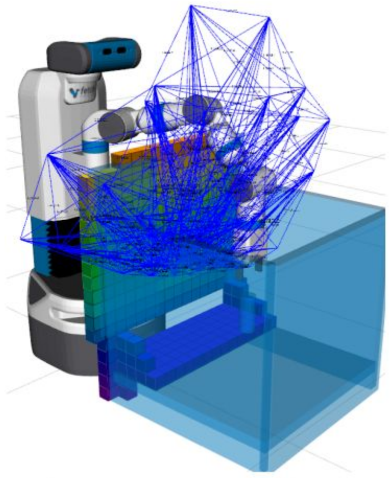
Evaluation: Baselines

Table 2. Baselines and BLIND

	Uses roadmap G	Uses human critiques	Update rule
Plain-TrajOpt	-	-	New random initialization
Random-BLIND	Yes	-	Apply random reward
Penalized-BLIND	Yes	Yes	Add fixed reward
BLIND	Yes	Yes	Learn the reward

Evaluation: Motion planning problem generation

$(X,Y)_{box} \pm 10cm$, $\Theta_{box} \pm 15^\circ$, $(S,G) \pm 10cm$ → Generate 10 environments → Generate 10 trajectories with TrajOpt → Compose roadmap by end-effector poses from these 10 trajectories

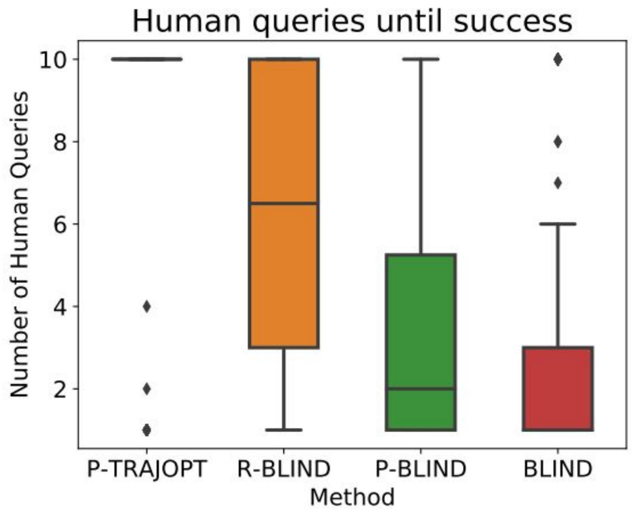


[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Evaluation: Results

AVG HUMAN QUERIES AND SUCCESS RATE

Method	Human Queries Mean (±std)	Success Rate
P-TRAJOPT	7.88 ± 3.78	0.24
R-BLIND	6.33 ± 3.05	0.67
P-BLIND	3.63 ± 3.19	0.87
BLIND	2.96 ± 3.17	0.86



[Motion Planning via Bayesian Learning in the Dark - Carlos Quintero-Pena, Constantinos Chamzas, Vaibhav Unhelkar and Lydia E. Kavraki, 2021]

Conclusion

- Producing safe trajectories with human guidance in incomplete environments.
- Only need a few interactions, even when environment is missing large parts.
- Applying BLIND to more varied environments and learning more general reward features.

References

- [1] J. Burlet, O. Aycard, and T. Fraichard, “Robust motion planning using markov decision processes and quadtree decomposition,” in IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA ’04. 2004, vol. 3, pp. 2820–2825 Vol.3, 2004.
- [2] J. van den Berg, S. Patil, and R. A. and, “Motion planning under uncertainty using iterative local optimization in belief space,” 2012.
- [3] J. Van Den Berg, P. Abbeel, and K. Goldberg, “LQG-MP: Optimized Path Planning for Robots with Motion Uncertainty and Imperfect State Information,” Int. J. Rob. Res., vol. 30, p. 895–913, June 2011.
- [4] C. Quintero-Pena , A. Kyrillidis, and L. E. Kavraki, “Robust optimization-based motion planning for high-DOF robots under sensing uncertainty,” in 2021 IEEE International Conference on Robotics and Automation (ICRA), 2021.
- [5] B. Saund and D. Berenson, “Motion planning for manipulators in unknown environments with contact sensing uncertainty,” in International Symposium on Robotics Research (ISRR), November 2018.
- [6] M. Palan, G. Shevchuk, N. Charles Landolfi, and D. Sadigh, “Learning Reward Functions by Integrating Human Demonstrations and Preferences,” in Robotics: Science and Syst., 2019.
- [7] Quintero-Pena, C., Chamzas, C., Unhelkar, V. and Kavraki, L.E., 2021, June. Motion planning via bayesian learning in the dark. In ICRA*:* Workshop on Machine Learning for Motion Planning.

Thanks for your attention