Monte-Carlo Planning in Hybrid Belief POMDPs

Moran Barenboim, Moshe Shienman and Vadim Indelman

IEEE Robotics and Automation Letters (RA-L), 8(8): 4410-4417, 2023.

TECHNION
Israel Institute of Technology

ICRA2024
May 13-17, Yokohama, Japan

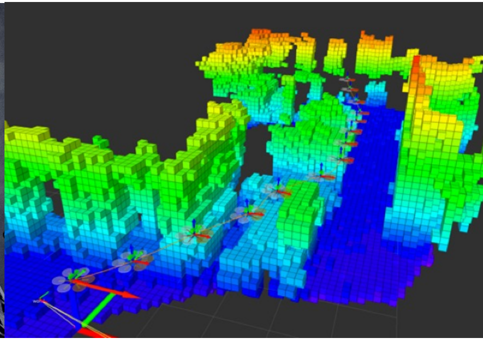
ANPL
Autonomous Navigation and Perception Lab

May 17, 2024

Introduction

- Sequential decision-making under uncertainty
- These are commonly formalized as POMDPs



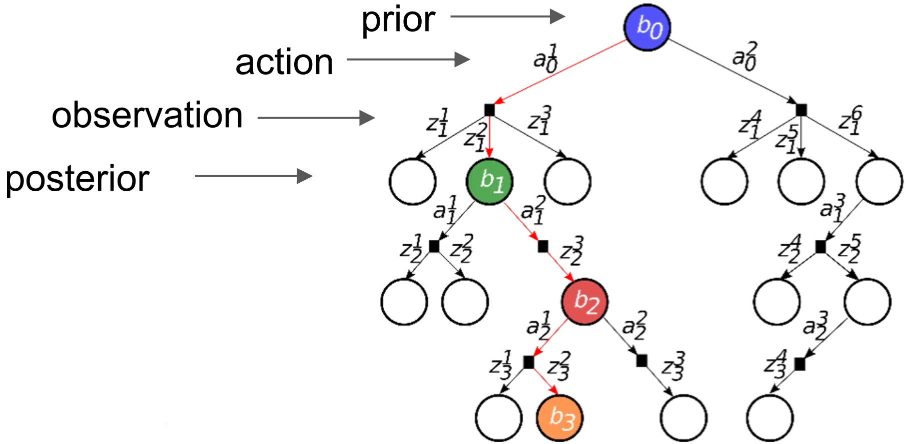



Introduction

A common approach for solving POMDPs is through online tree search methods

- Each node represents a belief
- Each edge represents an action or an observation
- Given a prior belief, the posterior belief is calculated via probabilistic inference

$$b_{t+1} = \psi(b_t, a_t, z_{t+1})$$



Autonomous Navigation and Perception Lab

Hybrid State Spaces

We will be focusing on POMDPs with Hybrid state spaces.

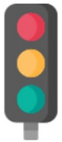
Continuous representations may include,

- The agent pose
- Landmark positio
- etc.

We name each realization of the discrete variables a hypothesis

While discrete variables may include,

- Object classes
- Data association hypotheses
- Semantic information (e.g. traffic light state)
- etc.



Hybrid State Spaces

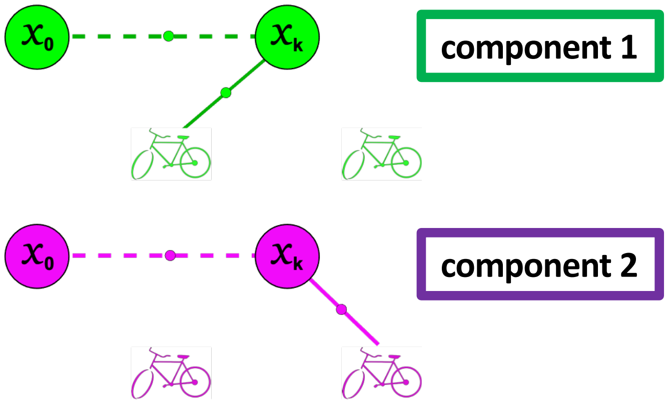
- Hybrid belief over continuous and discrete variables:

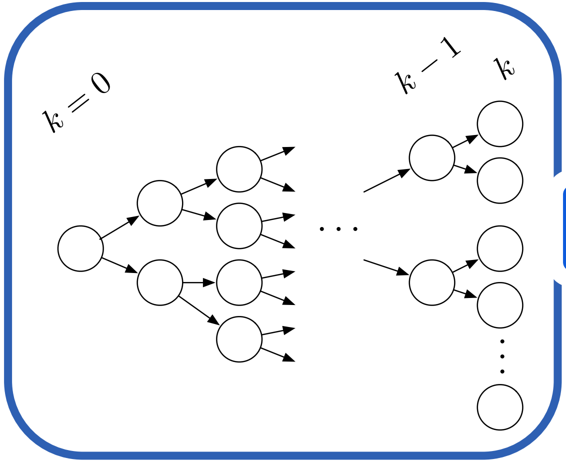
(e.g. agent state, landmark locations)

(e.g. data association hypotheses)

$$b_t \triangleq \mathbb{P}(X_t, \beta_{0:t} \mid H_t) = \underbrace{\mathbb{P}(X_t \mid \beta_{0:t}, H_t)}_{b[X_t]_{\beta_{0:t}}} \underbrace{\mathbb{P}(\beta_{0:t} \mid H_t)}_{b[\beta_{0:t}] \equiv \omega_t}.$$

- Belief over agent state is represented by a mixture density (e.g. GMM):



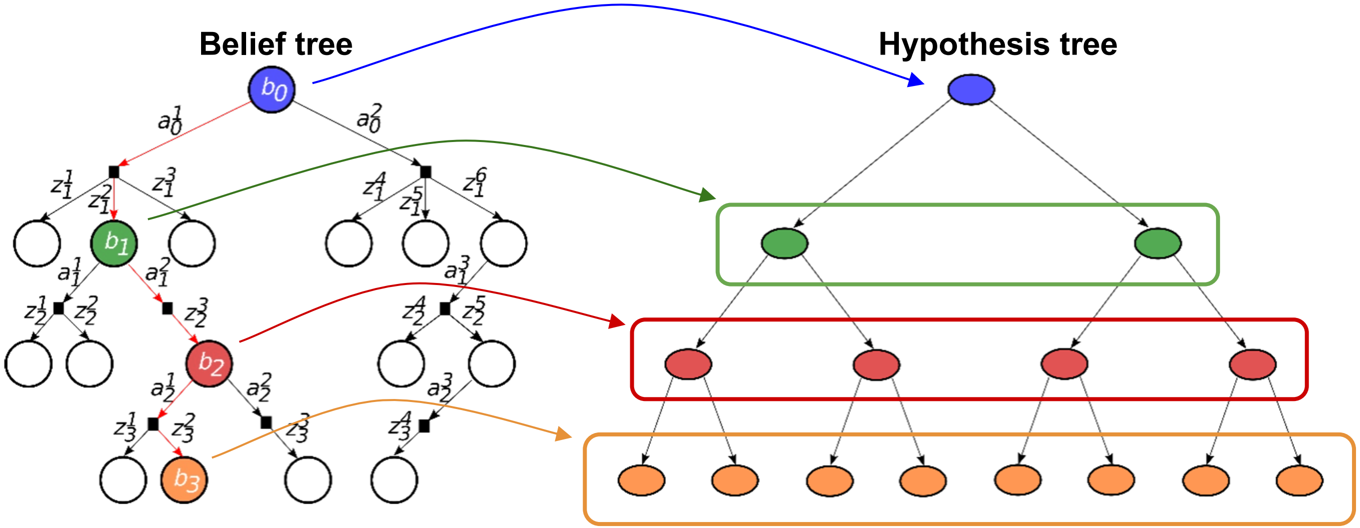


Exponential Growth

Hybrid State Spaces - The Challenge

Computing the value function requires explicit knowledge of the hypotheses

However, the number of hypotheses may grow **exponentially** with the horizon!

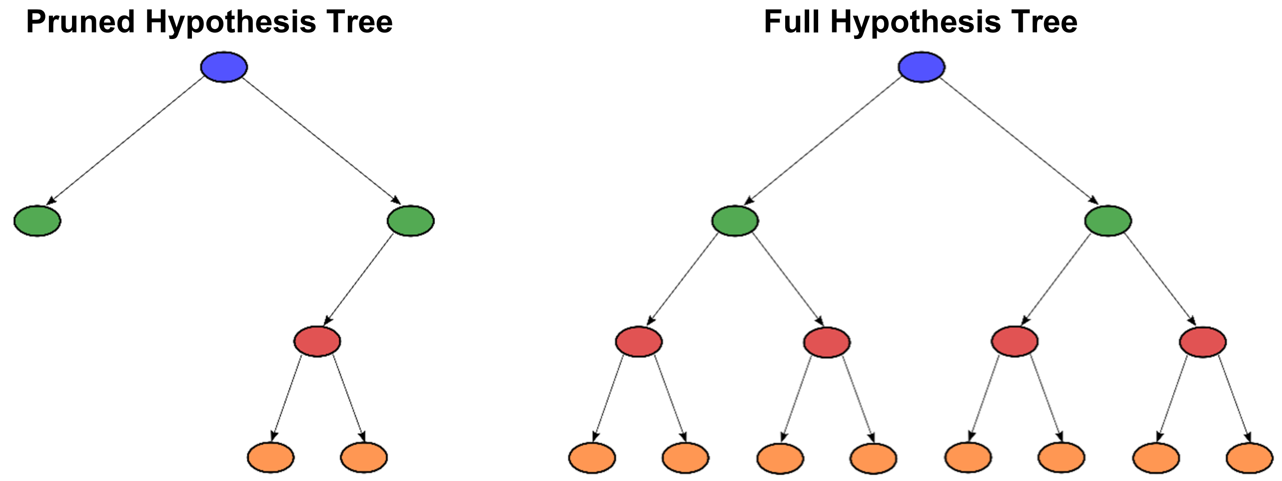


TECHNION Israel Institute of Technology
ANPL Autonomous Navigation and Perception Lab

Hybrid State Spaces - Current State Of The Art

Most SOTA algorithms prune hypotheses heuristically

However, this leads to a **biased estimator** of the value function



TECHNION Israel Institute of Technology

ANPL | Autonomous Navigation and Perception Lab

Hybrid State Spaces - Our Contributions

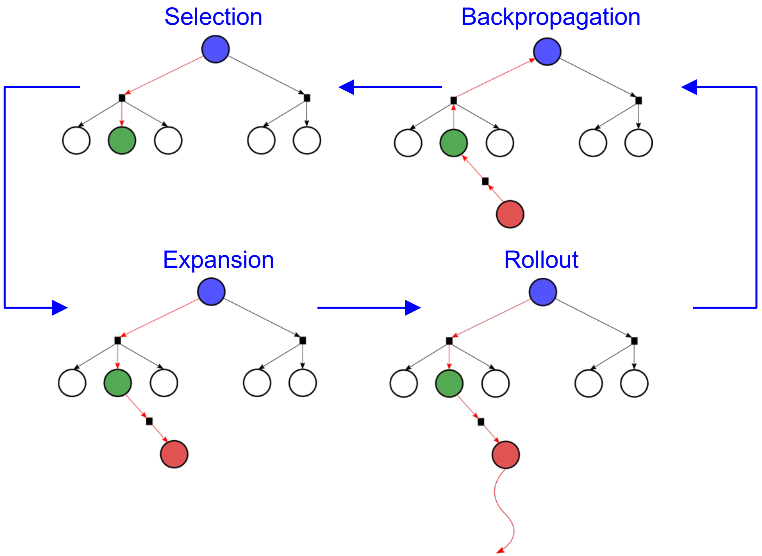
Instead of computing all possible hypotheses, we utilize MCTS sampling and exploration approach.

MCTS:

- An MDP solver
- Uses UCT to tradeoff exploration-exploitation for actions

$$UCT(x_t, a_t) = \hat{Q}(x_t, a_t) + c \cdot \sqrt{\frac{\log N(x_t)}{n(x_t, a_t)}}$$

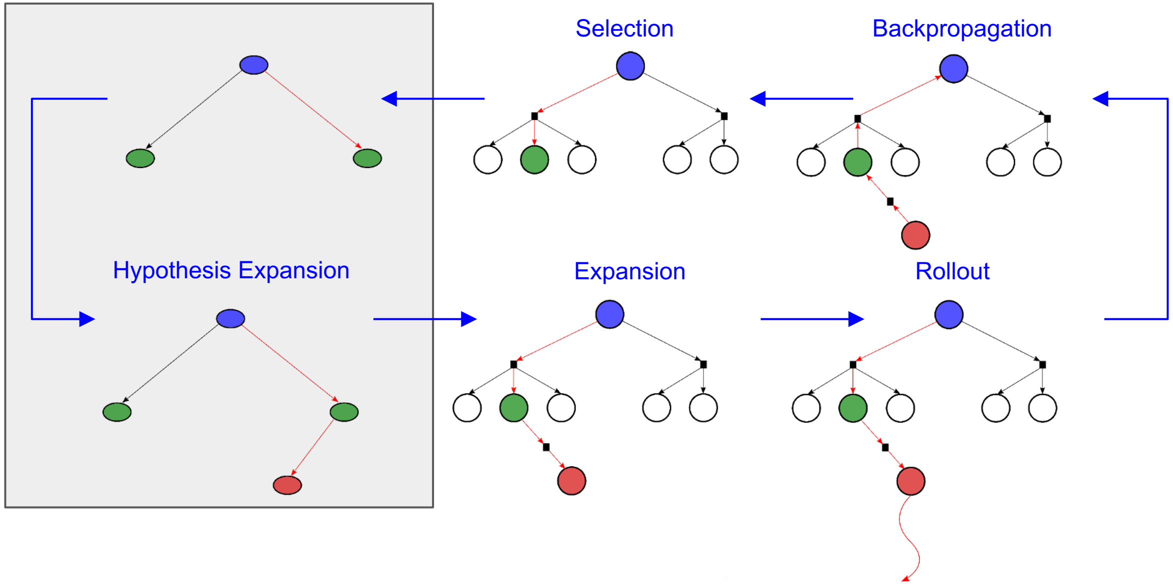
- Given an action, samples the next state
- Tends to quickly focus on the important parts of the tree



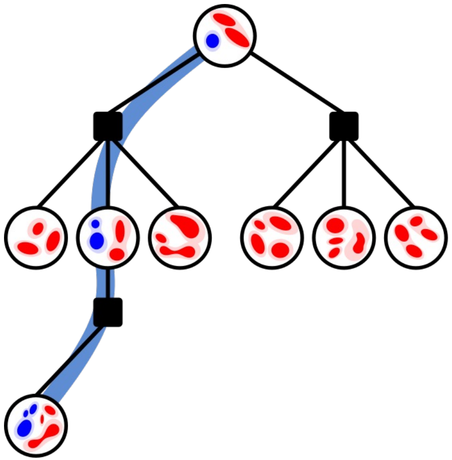
Hybrid State Spaces - Our Contributions

To solve Hybrid-POMDPs, we derived a new algorithm, named Hybrid-Belief Monte-Carlo Planning (HB-MCP)

HB-MCP adds a layer that samples hypotheses via Monte-Carlo sampling



Hybrid State Spaces - Our Contributions



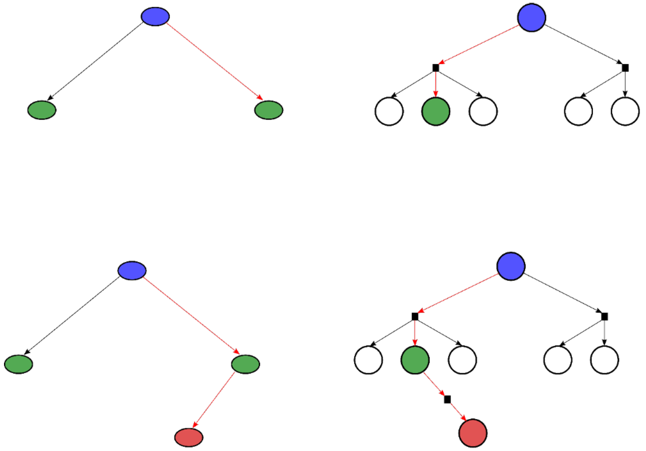
Full hybrid belief (shown in blue) at each iteration, regardless of the hypotheses significance.

Hybrid State Spaces - Our Contributions

We have derived a corresponding reward estimator, $\hat{\mathcal{R}}_X$, and have shown that it leads to an unbiased estimator,

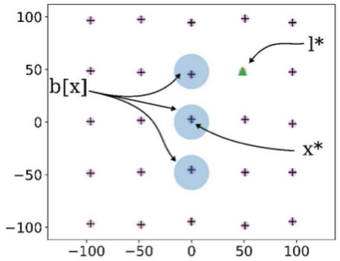
Lemma

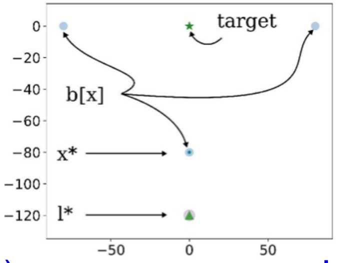
The sampled-based, state-dependent reward estimator, $\hat{\mathcal{R}}_X \triangleq \frac{1}{N}\sum_{i,j=1}^{N} \lambda_t^{i,j} \frac{1}{n_X}\sum_{k=1}^{n_X} r(X_t^{i,j,k}, a_t)$*, is unbiased.*

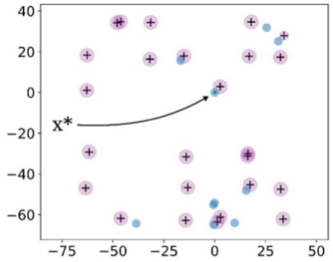


Hybrid State Spaces - Results

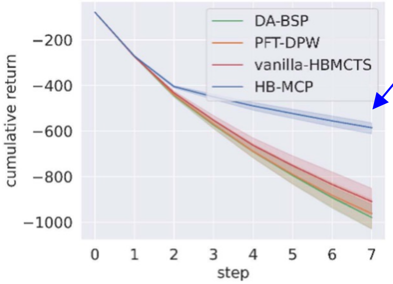
HB-MCP has shown improved performance on multiple experiments



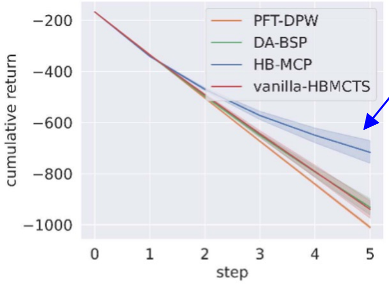




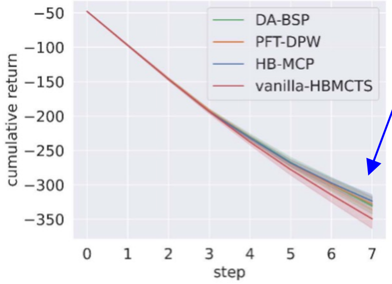
HB-MCP (ours)



HB-MCP (ours)



HB-MCP (ours)



Hybrid State Spaces - Conclusions

To conclude,

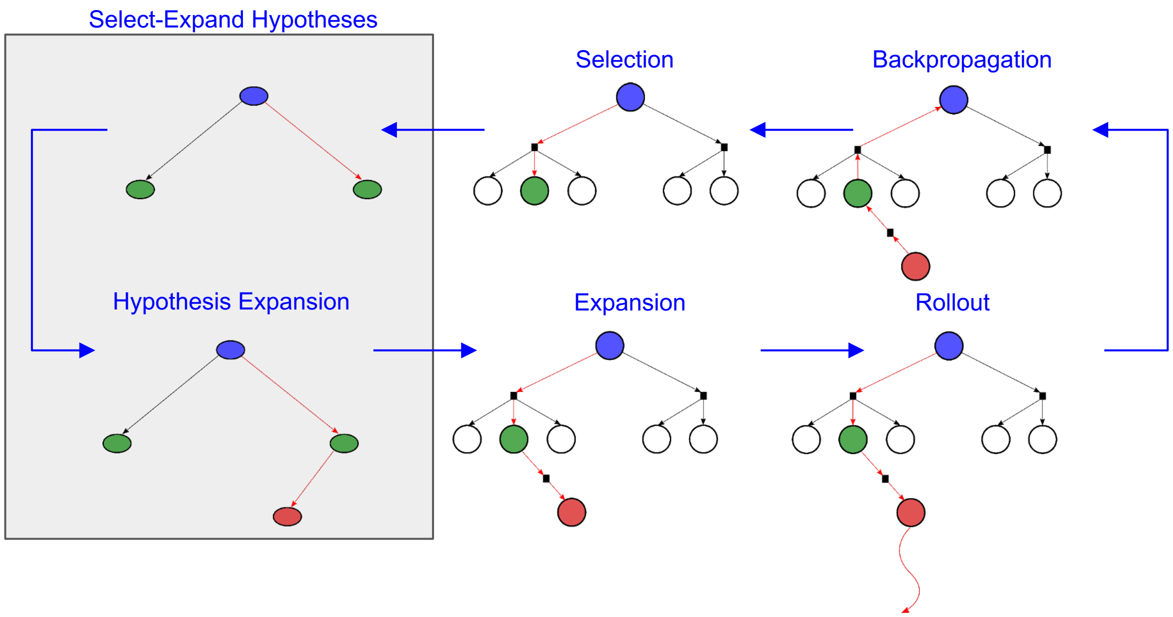
- Hybrid POMDPs are computationally difficult and the number of hypotheses may even grow exponentially with the horizon

- Naively pruning hypotheses leads to a biased estimation of the value function

- Instead, we suggest a new algorithm and show that it leads to an unbiased estimator

Monte-Carlo Planning in Hybrid Belief POMDPs

Moran Barenboim, Moshe Shienman and Vadim Indelman



TECHNION
Israel Institute
of Technology
ANPL
Autonomous Navigation
and Perception Lab