$N(s,a)$
$Q(s,a)$

Online POMDP Methods

Approximate POMDP Solutions

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

↓

Offline

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

↙ ↘

Offline **Online**

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

Offline

Previously

Online

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

Offline

Previously

Online

Formulation Approximations

(solve a slightly different problem)

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

Offline

Previously

Online

Formulation Approximations

(solve a slightly different problem)

Last Time

Approximate POMDP Solutions

Numerical Approximations

(approximately solve original problem)

Offline

Previously

Online

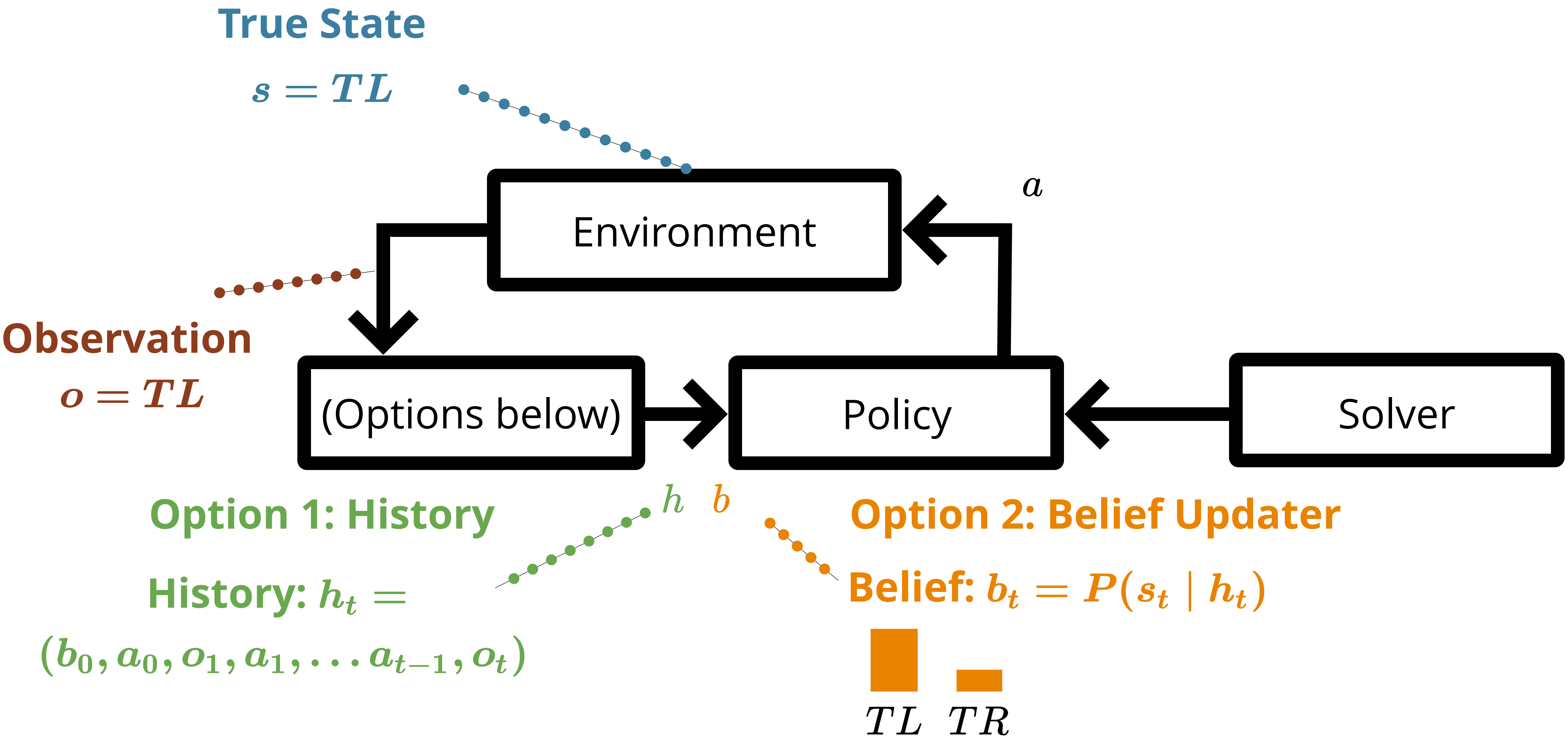
Today!

Formulation Approximations

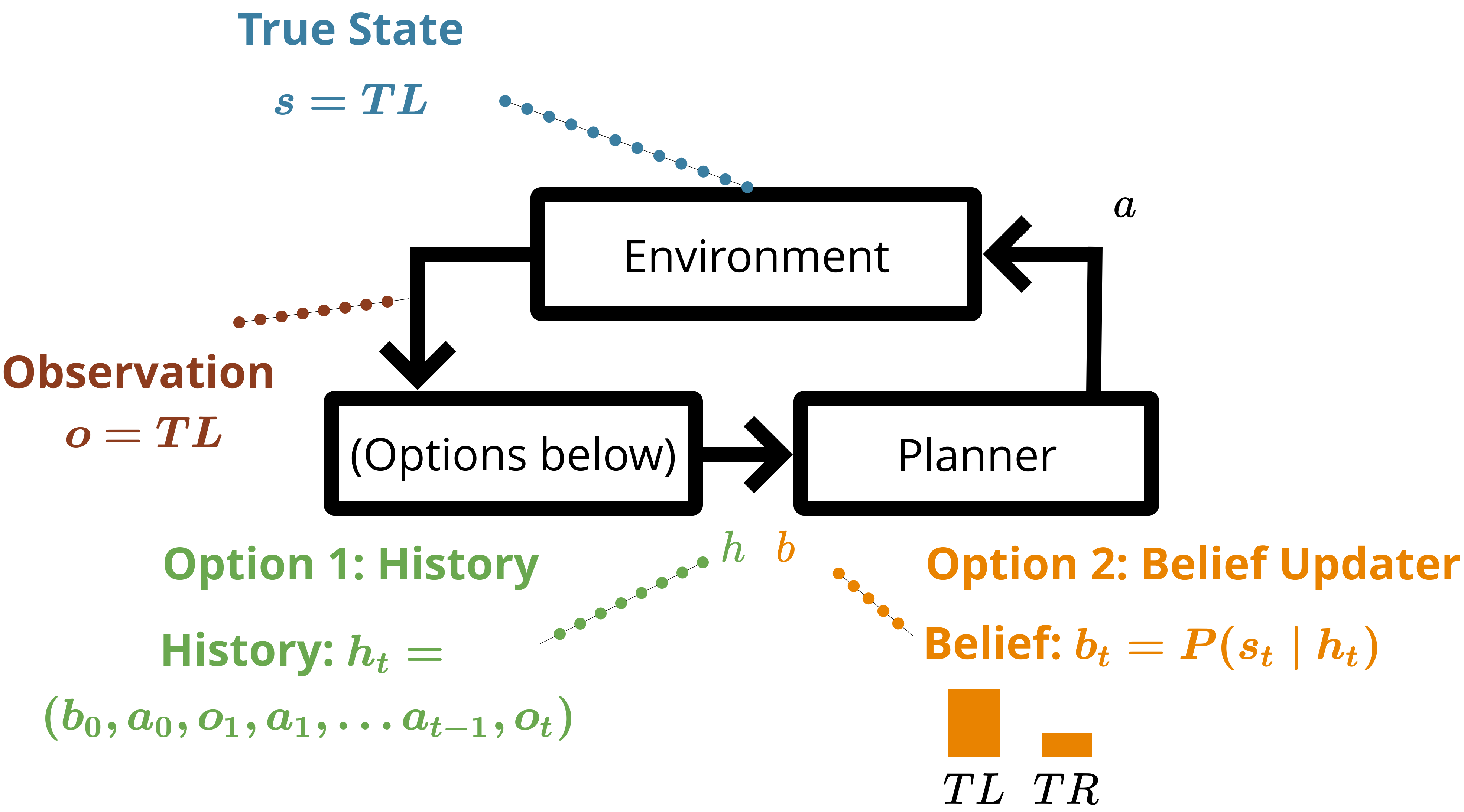
(solve a slightly different problem)

Last Time

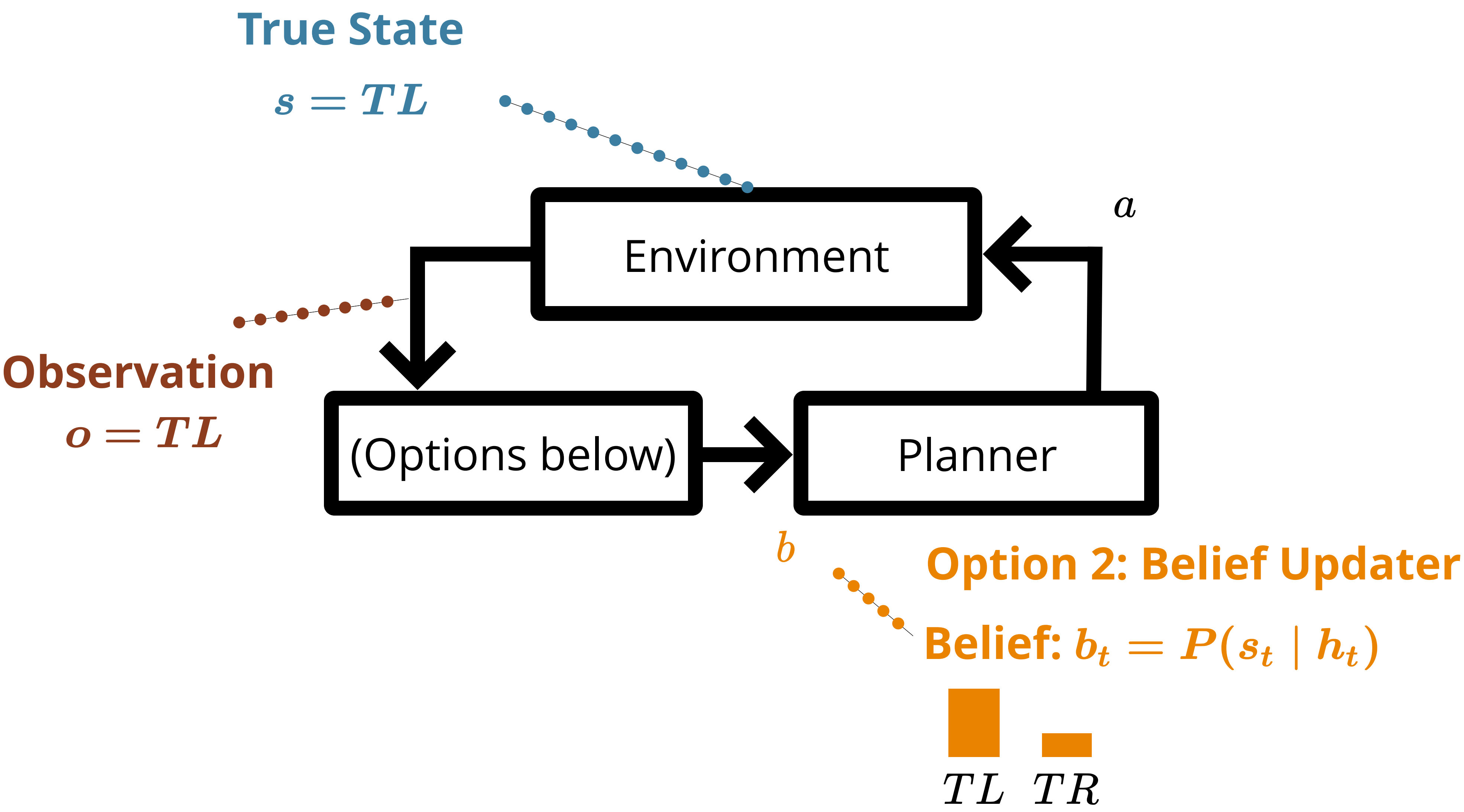
POMDP Sense-Plan-Act Loop



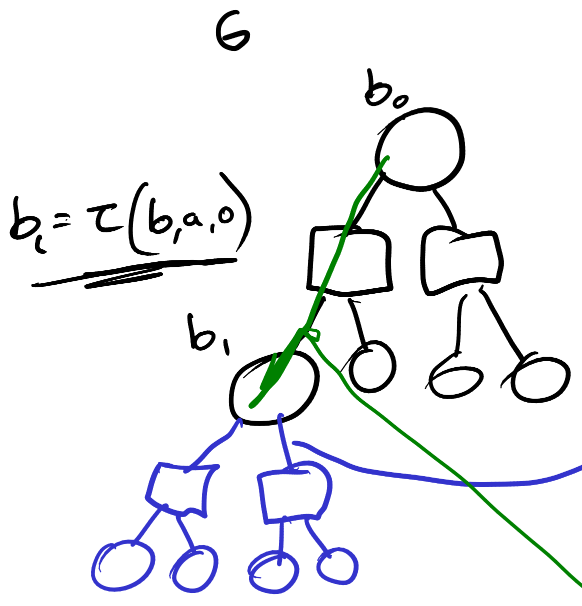
POMDP Sense-Plan-Act Loop



POMDP Sense-Plan-Act Loop



Belief-Space Tree Search: AEMS



while time remains

$$b^* = \operatorname*{argmax}_{b \in fringe(G)} E(b)$$

expand (b^*)

backup (b^*)

$$E(b) = \gamma^d P(b) \hat{\varepsilon}(b)$$

$$\hat{\varepsilon}(b) = U(b) - L(b)$$ ← domain-specific heuristic

$$P(b) = \prod_{i=0}^{d-1} P(o_i \mid b_i, a_i) \, P(a_i \mid b_i)$$

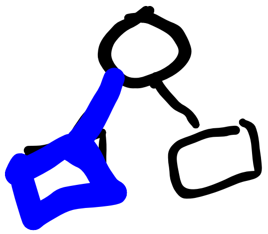
$$\sum_{s,s'} T(s' \mid s,a) \, Z(o \mid a,s') \, b(s)$$

$$P(a \mid b) \propto \frac{U(b,a) - L(b)}{U(b) - L(b)}$$ ← AEMS1

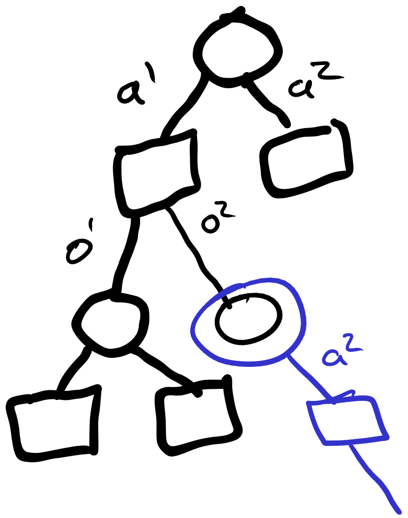
$$P(a \mid b) \propto \begin{cases} 1 & \text{if } a = \operatorname*{argmax}_{a'} U(b,a') \\ 0 & \text{o.w.} \end{cases}$$ ← AEMS2

Monte Carlo Tree Search (MCTS/UCT)

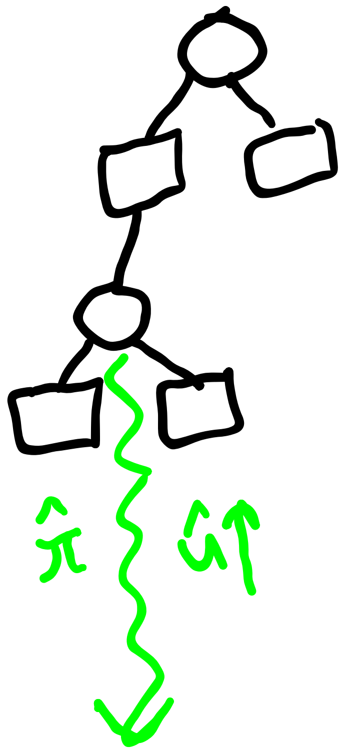
Search



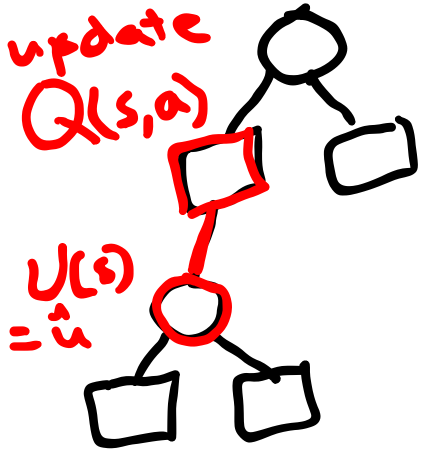
Expansion



Rollout



Backup

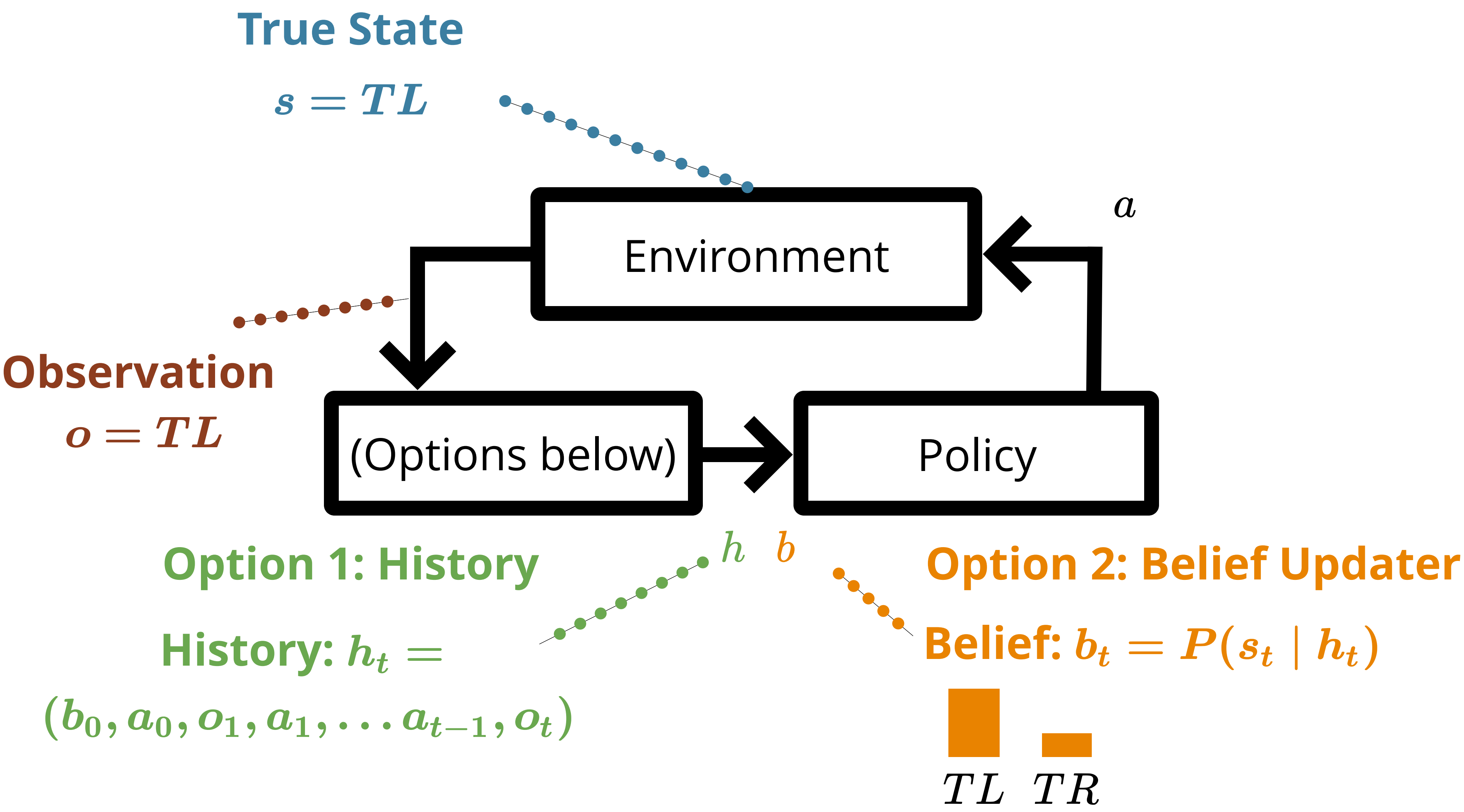


$$Q(s,a) + c\sqrt{\frac{\log N(s)}{N(s,a)}}$$

low $N(s,a)/N(s)$ = high bonus

start with $c = 2(\bar{V} - \underline{V})$

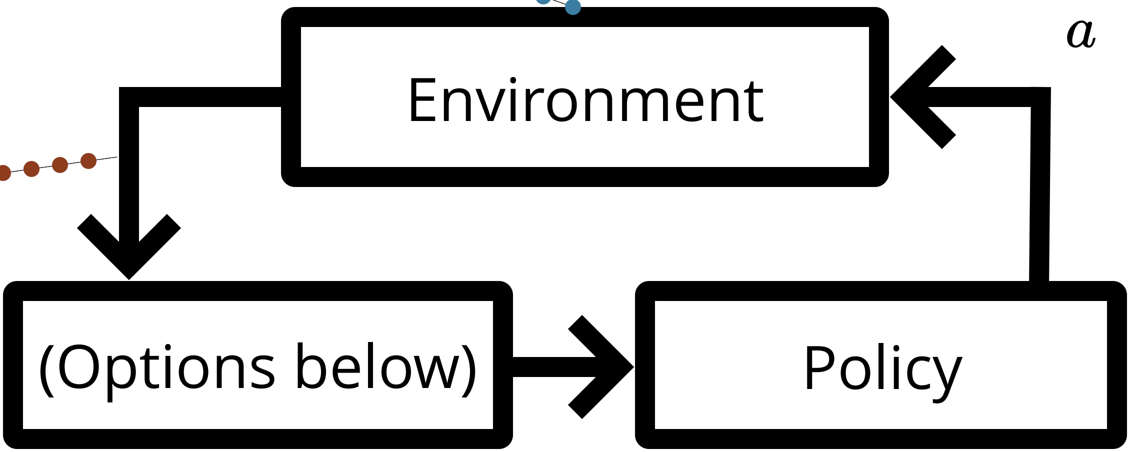
How should we adapt MCTS for POMDPs?



How should we adapt MCTS for POMDPs?

True State

$s = TL$



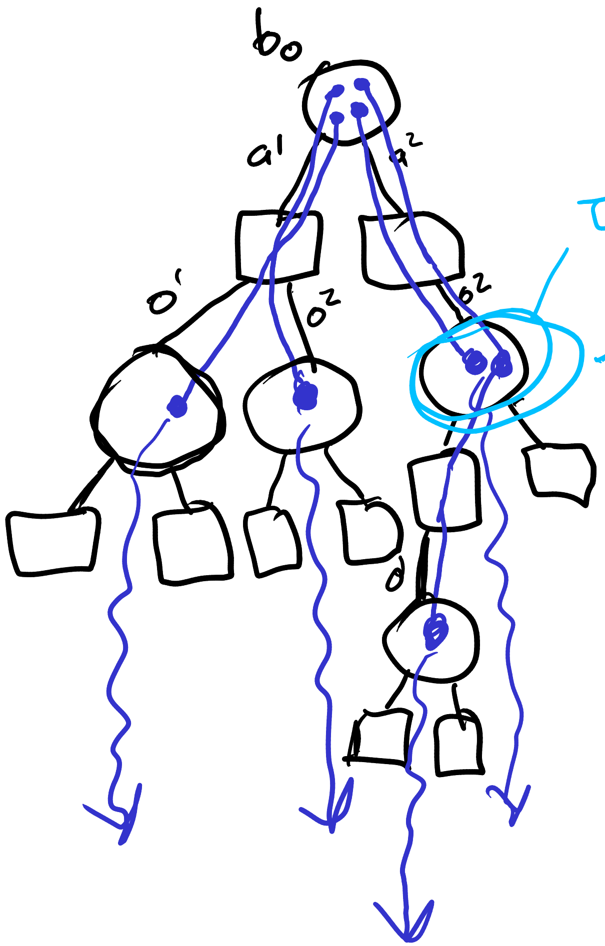
Observation

$o = TL$

Option 1: History

History: $h_t = (b_0, a_0, o_1, a_1, \ldots a_{t-1}, o_t)$

MCTS on Histories

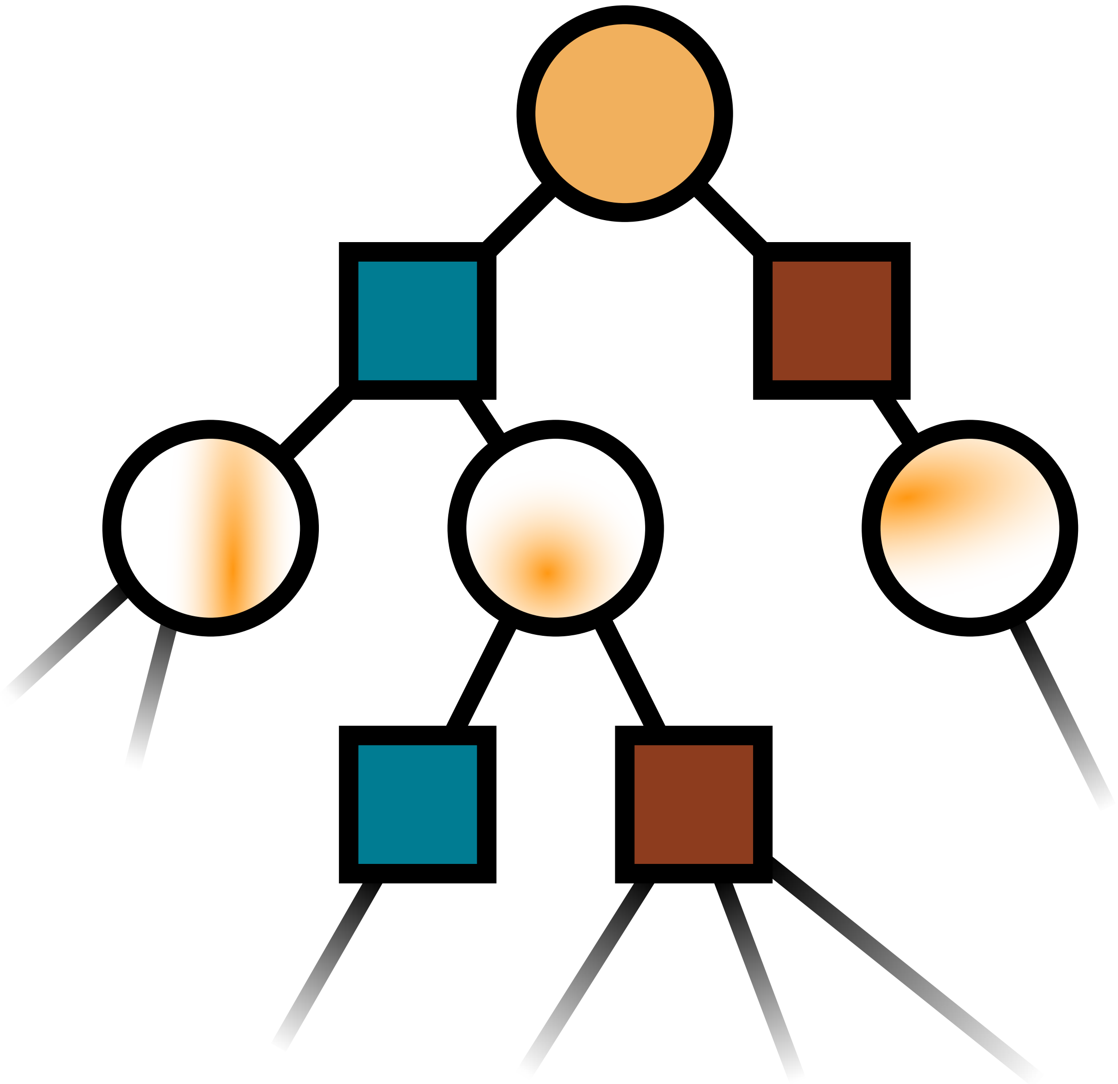


$\tau(b_0, a^2, o^2)$

PO-UCT

POMCP

re-uses the particles from the tree for beliefs

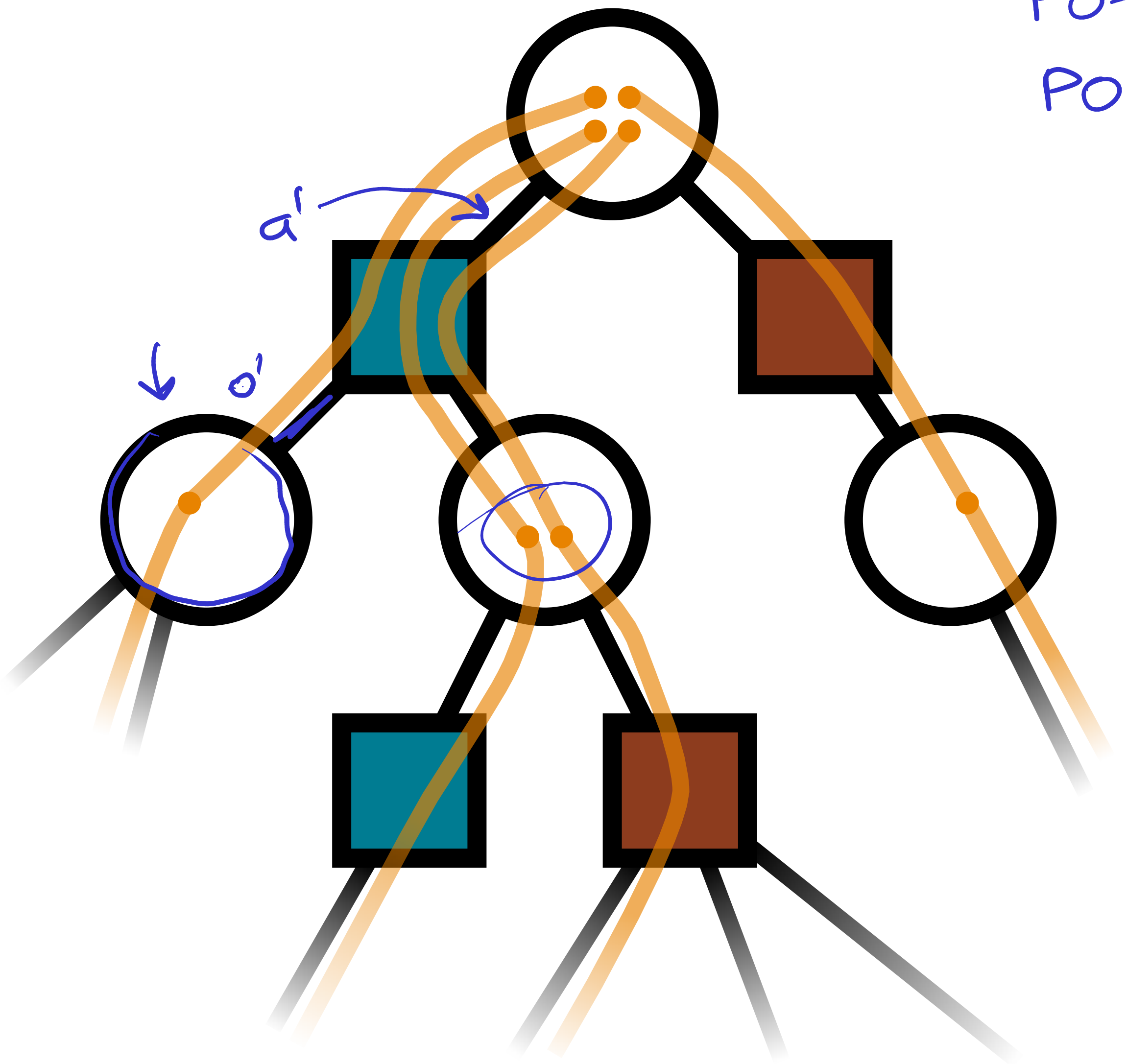


1. Online POMDP
2. Application
3. OOP
4. HW6

history MCTS

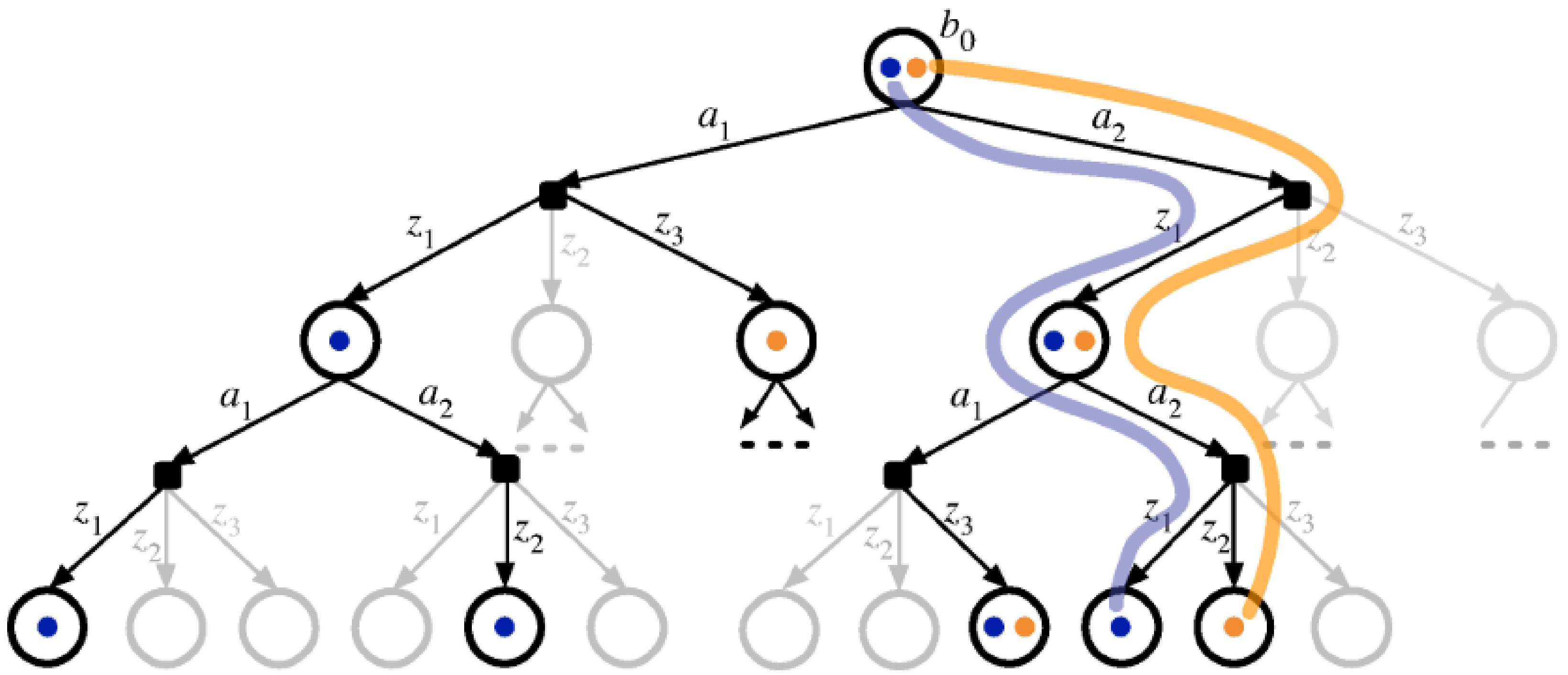
PO-UCT

POMCP



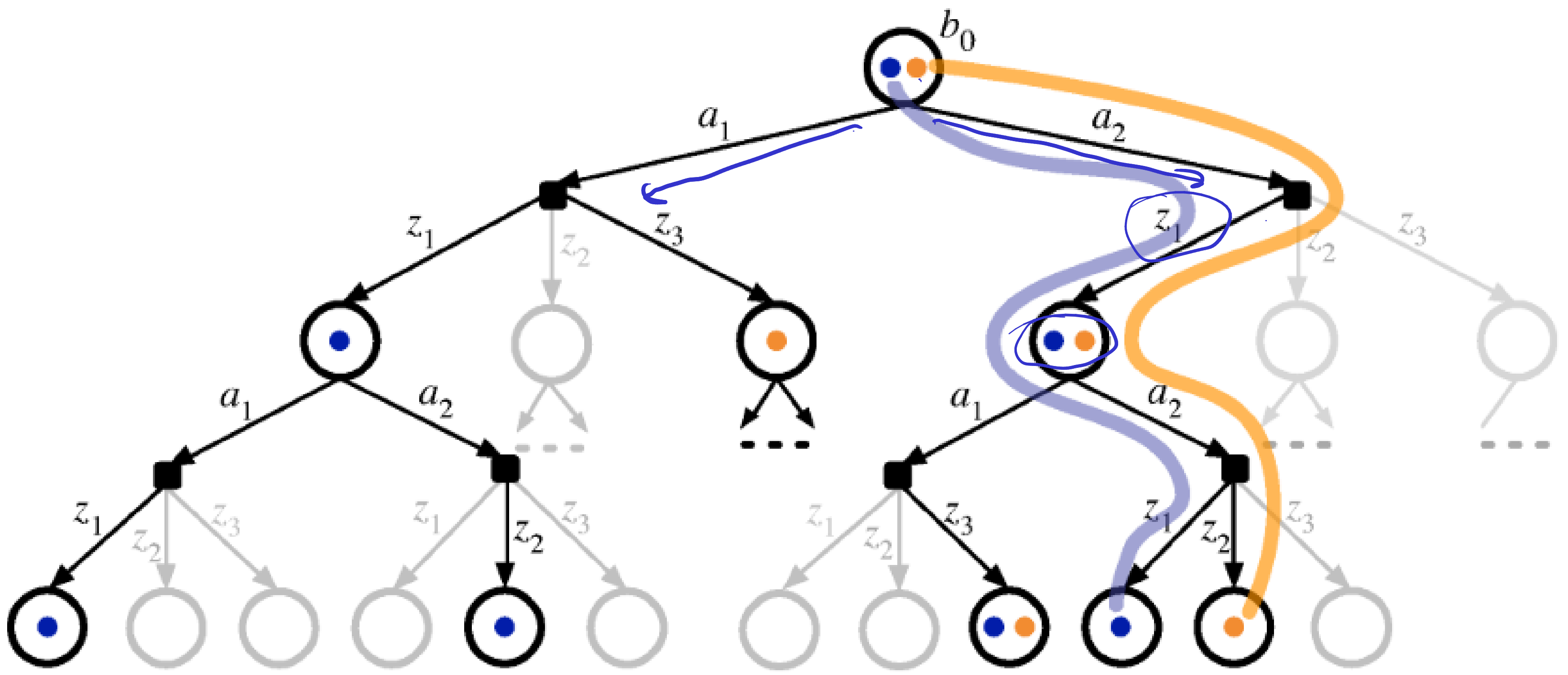
DESPOT

Determinized Sparse Partially Observable Tree



Somani, A., Ye, N., Hsu, D., & Lee, W. "DESPOT : Online POMDP Planning with Regularization." *Journal of Artificial Intelligence Research*, 2017

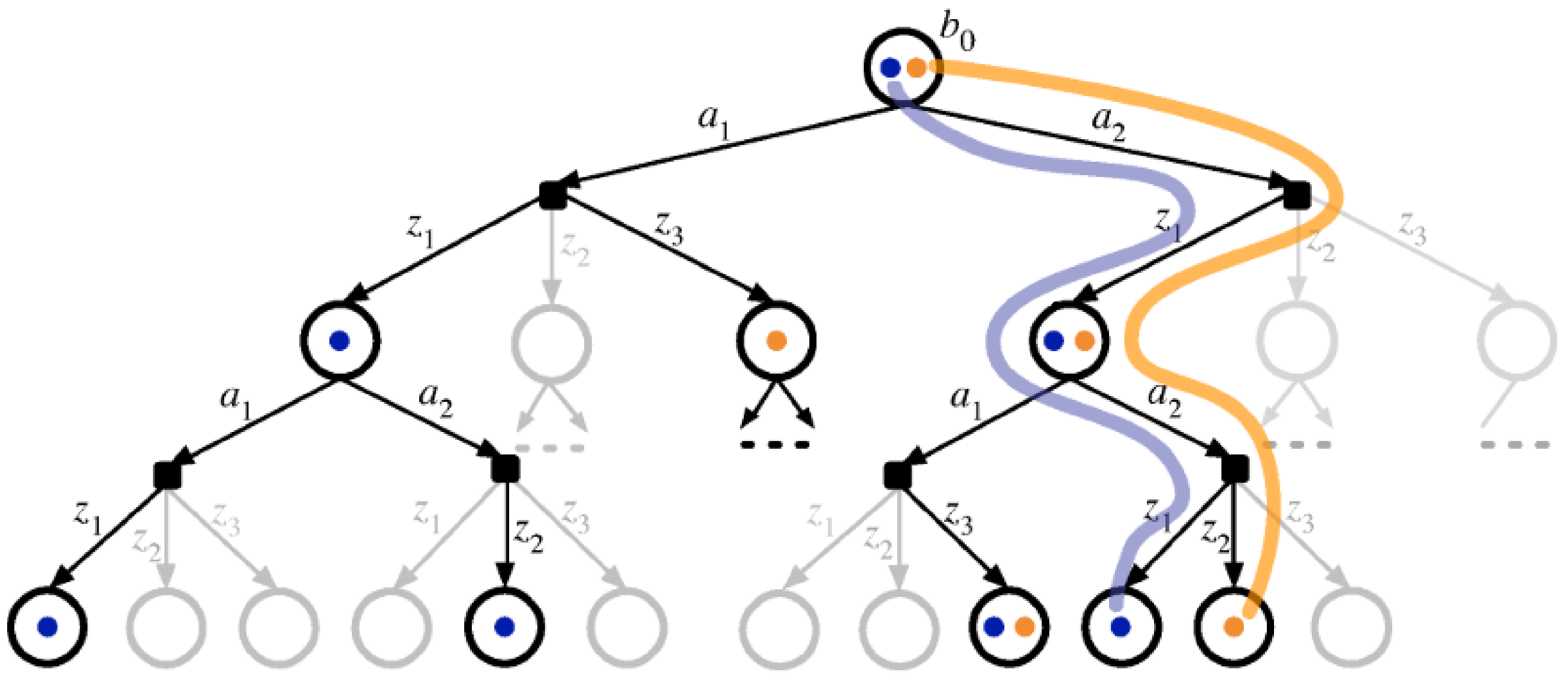
DESPOT



- Determinized Scenarios

Somani, A., Ye, N., Hsu, D., & Lee, W. "DESPOT : Online POMDP Planning with Regularization." *Journal of Artificial Intelligence Research*, 2017

DESPOT

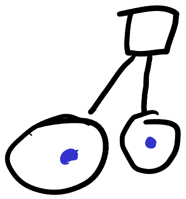


$\varepsilon(b) = U(b) - L(b)$

$U(b)$

$L(b)$

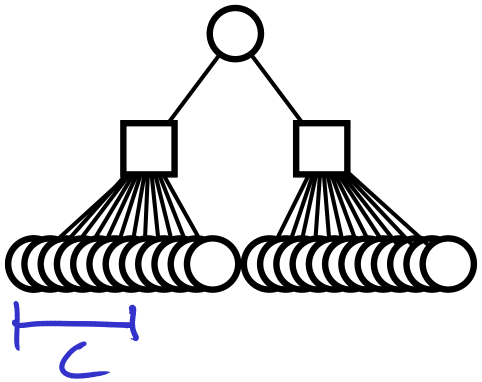
- Determinized Scenarios
- Guided by Lower and Upper Bounds

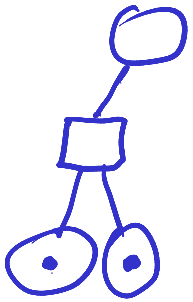


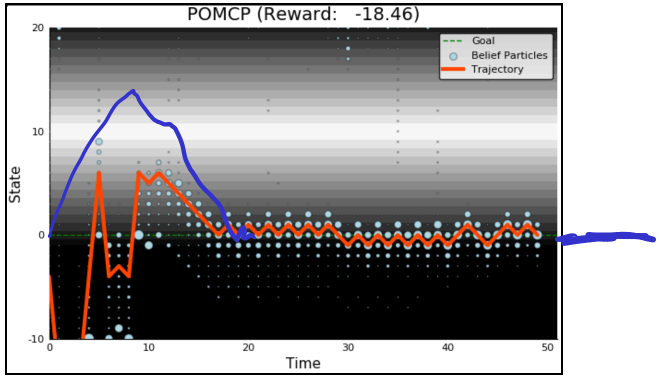
Somani, A., Ye, N., Hsu, D., & Lee, W. "DESPOT : Online POMDP Planning with Regularization." *Journal of Artificial Intelligence Research*, 2017

Continuous Observation Spaces

POMCP



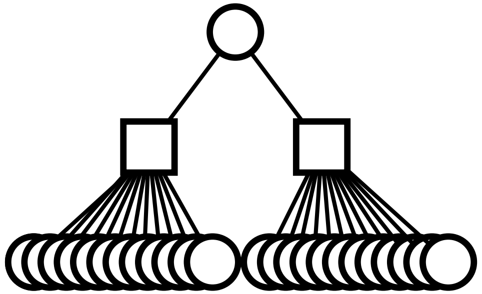


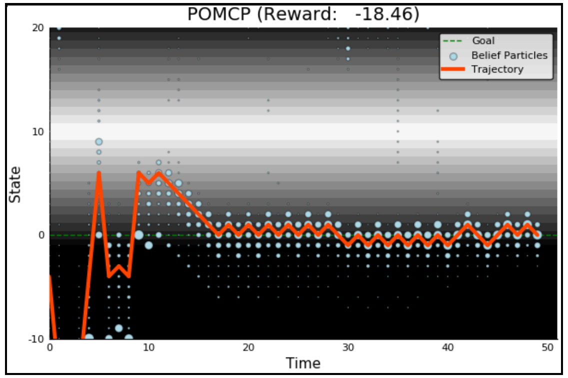


Continuous Observation Spaces

POMCP

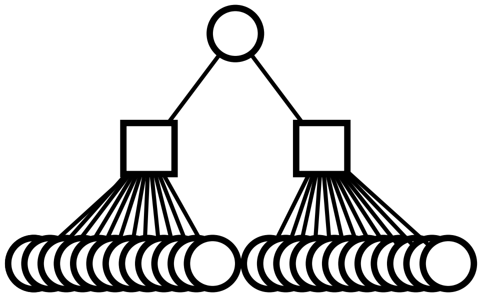
POMCPOW

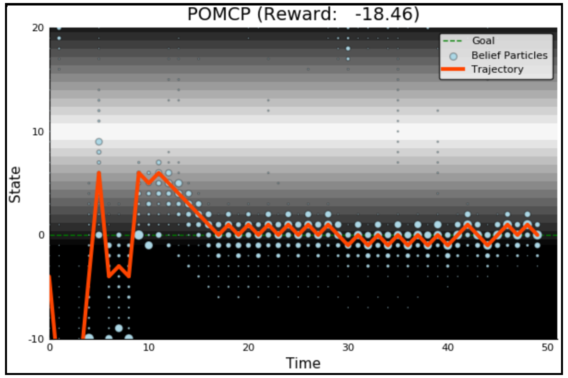




Continuous Observation Spaces

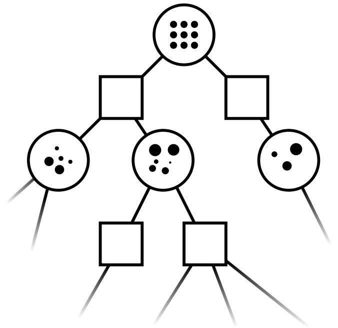
POMCP





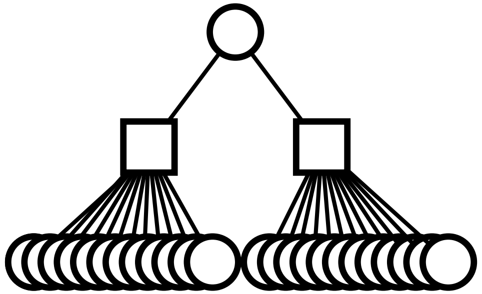
POMCPOW

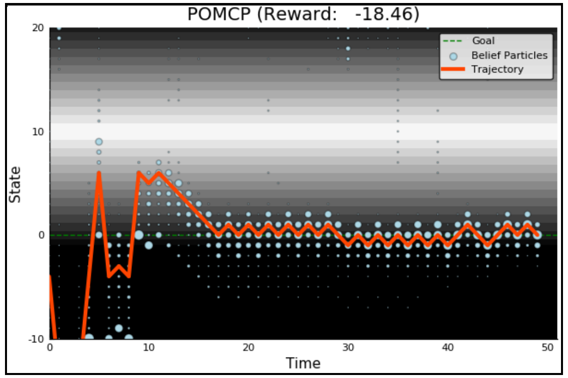
DPW



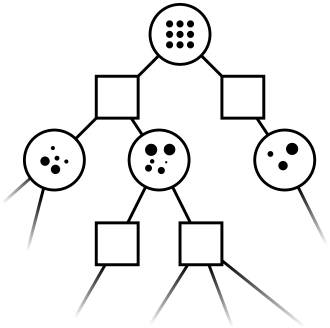
Continuous Observation Spaces

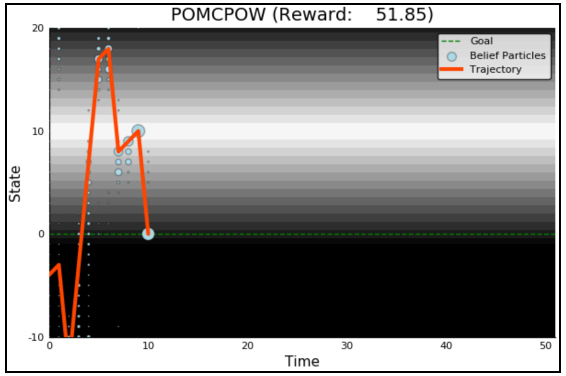
POMCP





POMCPOW





PF Approximation Accuracy

[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

PF Approximation Accuracy

$b(s) \approx \sum_i w_i \mathbb{1}(s_i = s)$ $\mathbf{M_P}$ = Particle belief MDP approximation of POMDP $\mathbf{P}$

[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

PF Approximation Accuracy

$\mathbf{M_P}$ = Particle belief MDP approximation of POMDP $\mathbf{P}$

For any $\epsilon > 0$ and $\delta > 0$, if C (number of particles) is high enough,

[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

PF Approximation Accuracy

$\mathbf{M_P}$ = Particle belief MDP approximation of POMDP $\mathbf{P}$

For any $\epsilon > 0$ and $\delta > 0$, if C (number of particles) is high enough,

$$|Q^*_{\mathbf{P}}(b, a) - Q^*_{\mathbf{M_P}}(\bar{b}, a)| \leq \epsilon \quad \text{w.p. } 1 - \delta$$

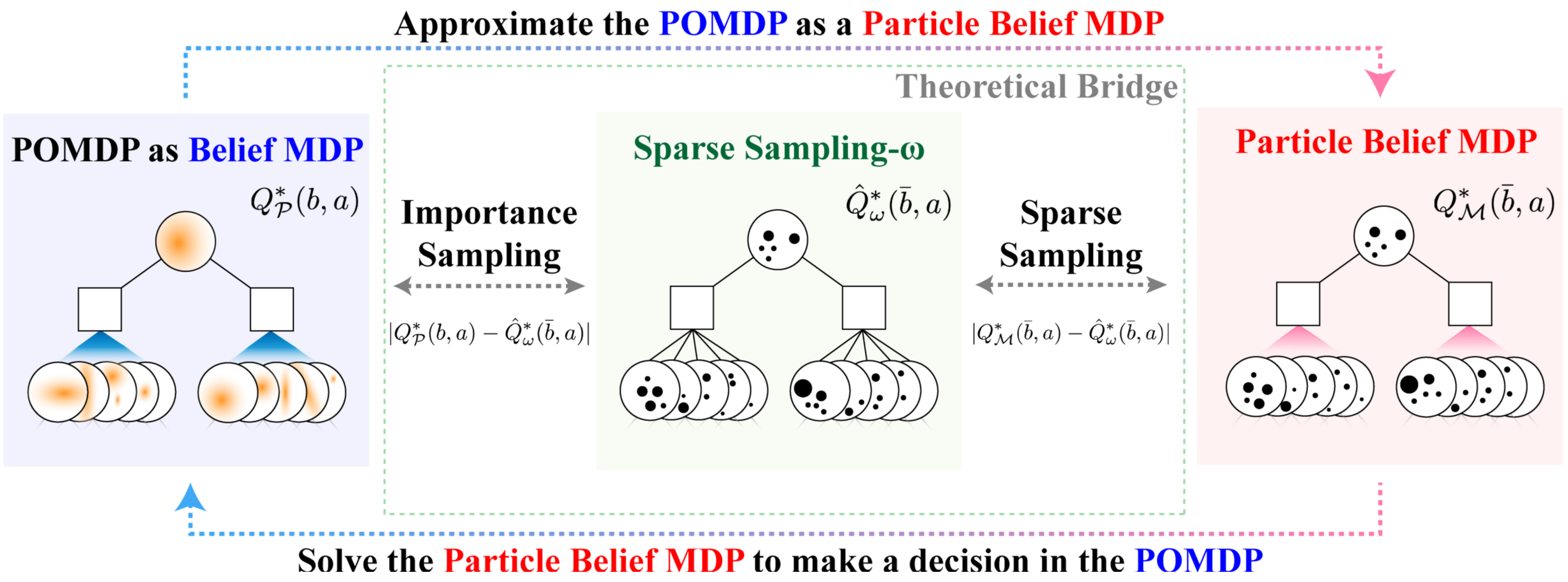
[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

PF Approximation Accuracy

$\mathbf{M_P}$ = Particle belief MDP approximation of POMDP $\mathbf{P}$

For any $\epsilon > 0$ and $\delta > 0$, if C (number of particles) is high enough,

$$|Q^*_{\mathbf{P}}(b,a) - Q^*_{\mathbf{M_P}}(\bar{b},a)| \leq \epsilon \quad \text{w.p. } 1-\delta$$



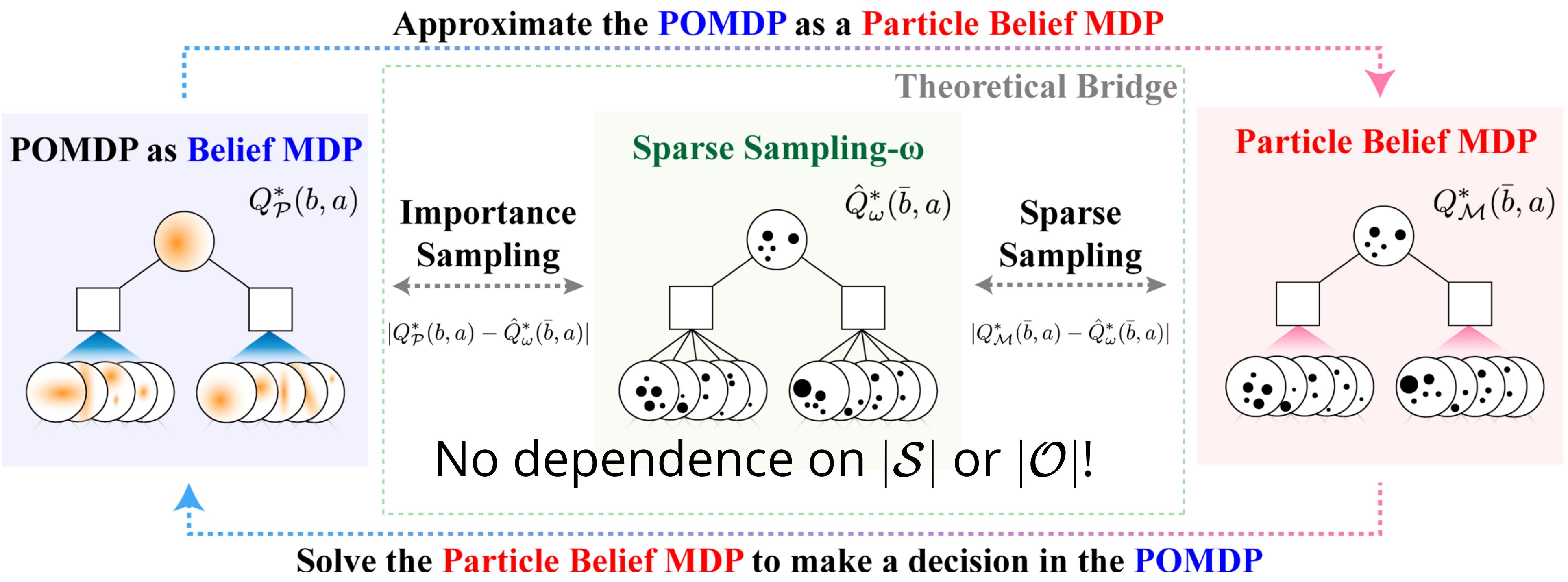
[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

PF Approximation Accuracy

$\mathbf{M_P}$ = Particle belief MDP approximation of POMDP $\mathbf{P}$

For any $\epsilon > 0$ and $\delta > 0$, if C (number of particles) is high enough,

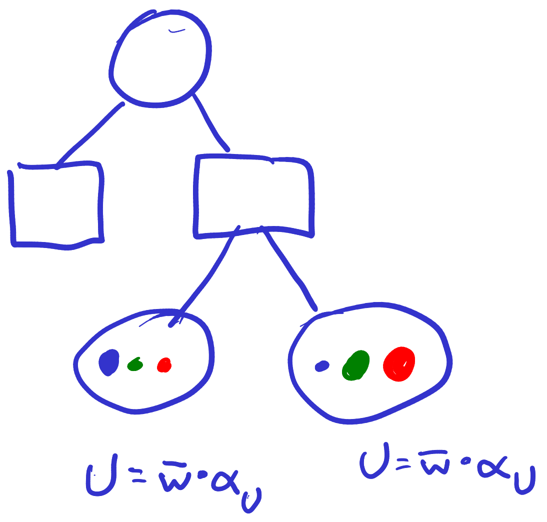
$$|Q^*_{\mathbf{P}}(b,a) - Q^*_{\mathbf{M_P}}(\bar{b},a)| \leq \epsilon \quad \text{w.p. } 1-\delta$$



No dependence on $|\mathcal{S}|$ or $|\mathcal{O}|$!

[Lim, Becker, Kochenderfer, Tomlin, & Sunberg, 2023]

DESPOT-α

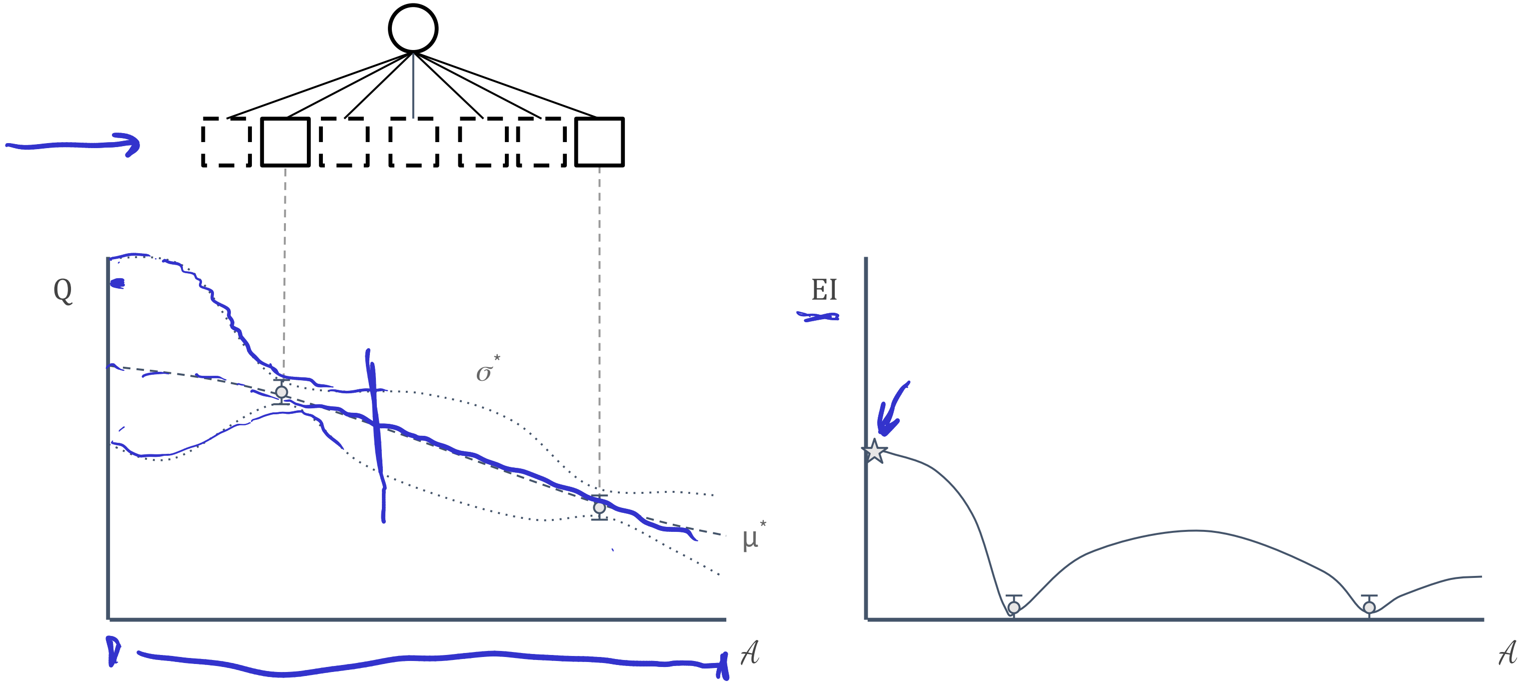


Continuous Action Spaces

Continuous Action Spaces

BOMCP

Bayesian Optimized Action Branching



Continuous Action Spaces

BOMCP

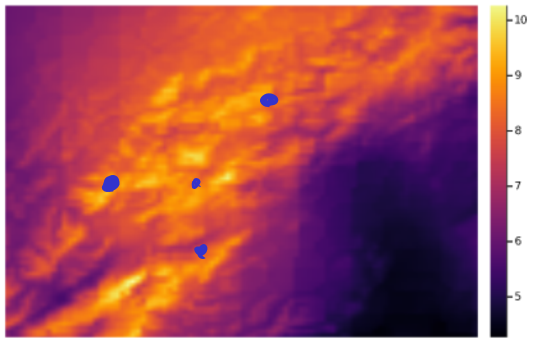
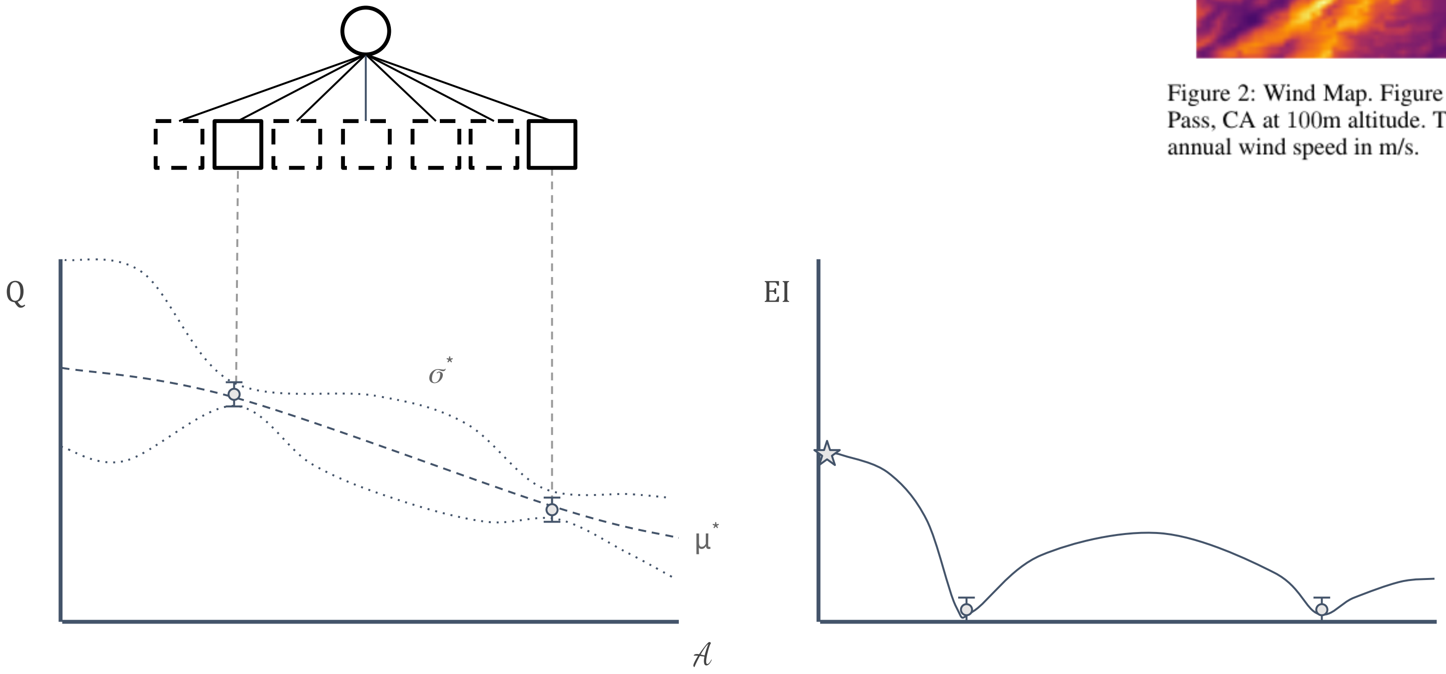


Figure 2: Wind Map. Figure shows wind map for Altamont Pass, CA at 100m altitude. The colors represent the average annual wind speed in m/s.

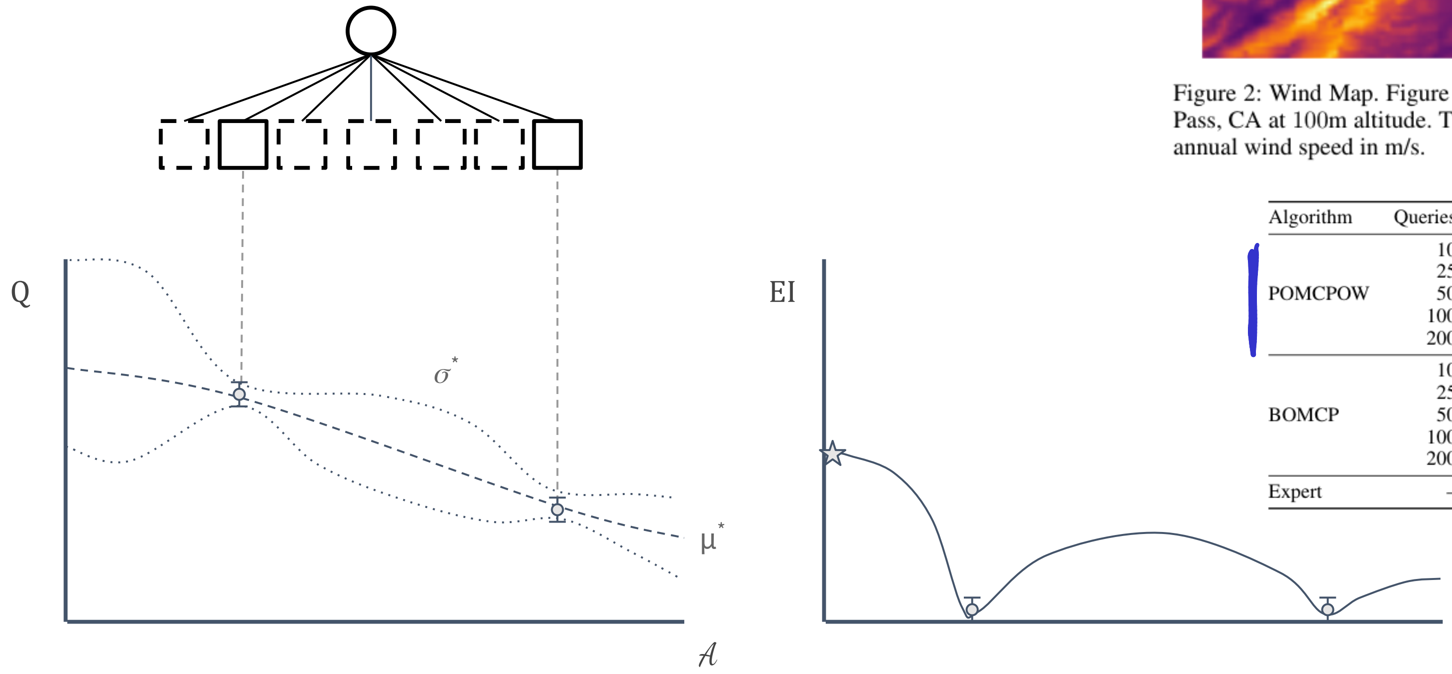
Bayesian Optimized Action Branching



Continuous Action Spaces

BOMCP

Bayesian Optimized Action Branching



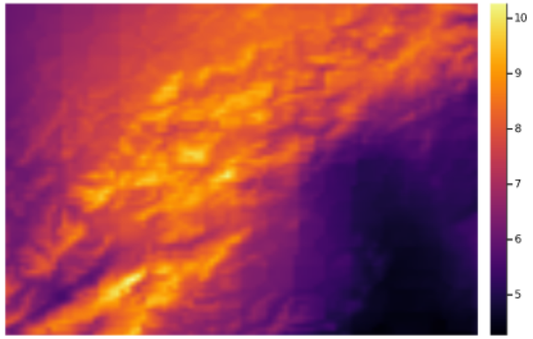


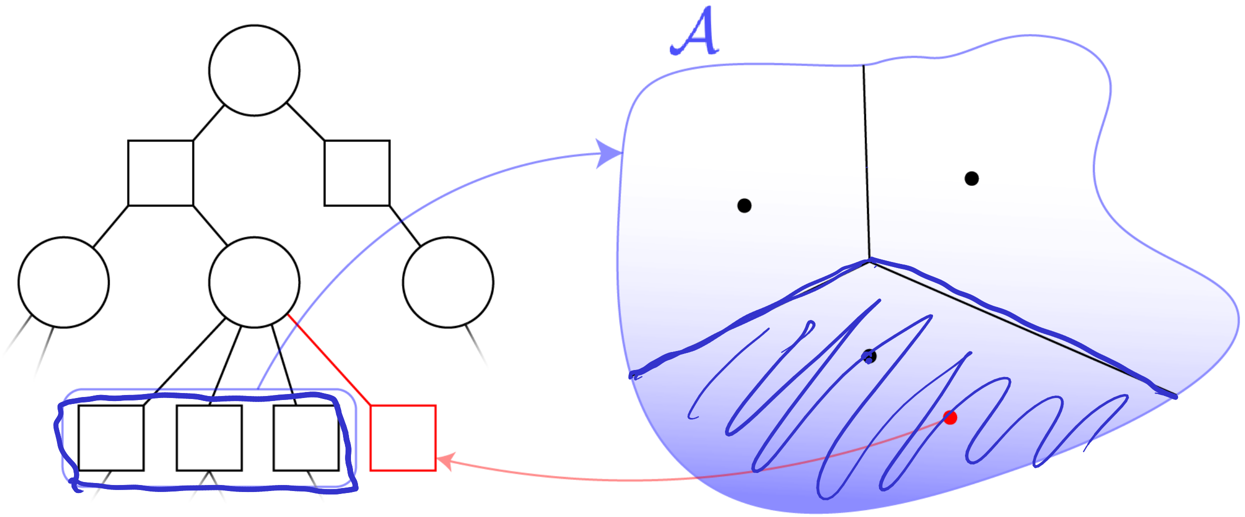
Figure 2: Wind Map. Figure shows wind map for Altamont Pass, CA at 100m altitude. The colors represent the average annual wind speed in m/s.

Algorithm	Queries	Score	Time (seconds)
POMCPOW	10	15708 ± 229	2.25 ± 0.07
	25	16234 ± 217	4.80 ± 0.07
	50	16374 ± 212	6.27 ± 0.08
	100	16018 ± 262	11.98 ± 0.07
	200	15787 ± 233	20.67 ± 0.09
BOMCP	10	18095 ± 183	2.55 ± 0.08
	25	18154 ± 158	5.21 ± 0.07
	50	18015 ± 163	6.71 ± 0.06
	100	18225 ± 119	13.39 ± 0.07
	200	18113 ± 157	25.14 ± 0.08
Expert	–	8130 ± 51	–

15000

18000

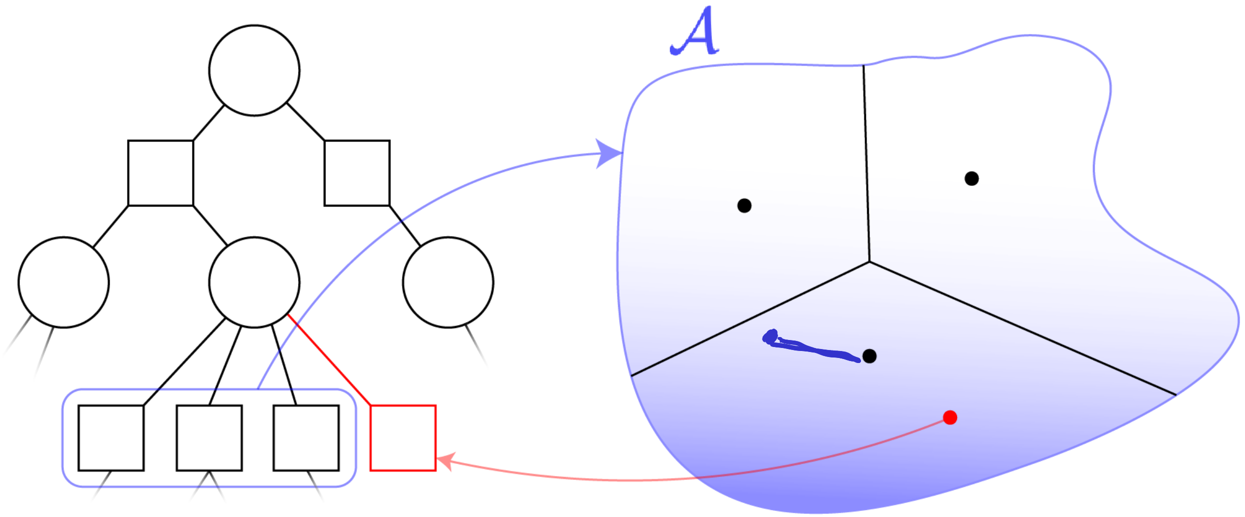
Voronoi Progressive Widening



Online Tree Search Planner

Voronoi Progressive Widening

Voronoi Progressive Widening



Online Tree Search Planner

Voronoi Progressive Widening