

# Online POMDP Methods

# Approximate POMDP Solutions

## Numerical Approximations

(approximately solve original problem)



**Offline**

Previously



**Online**

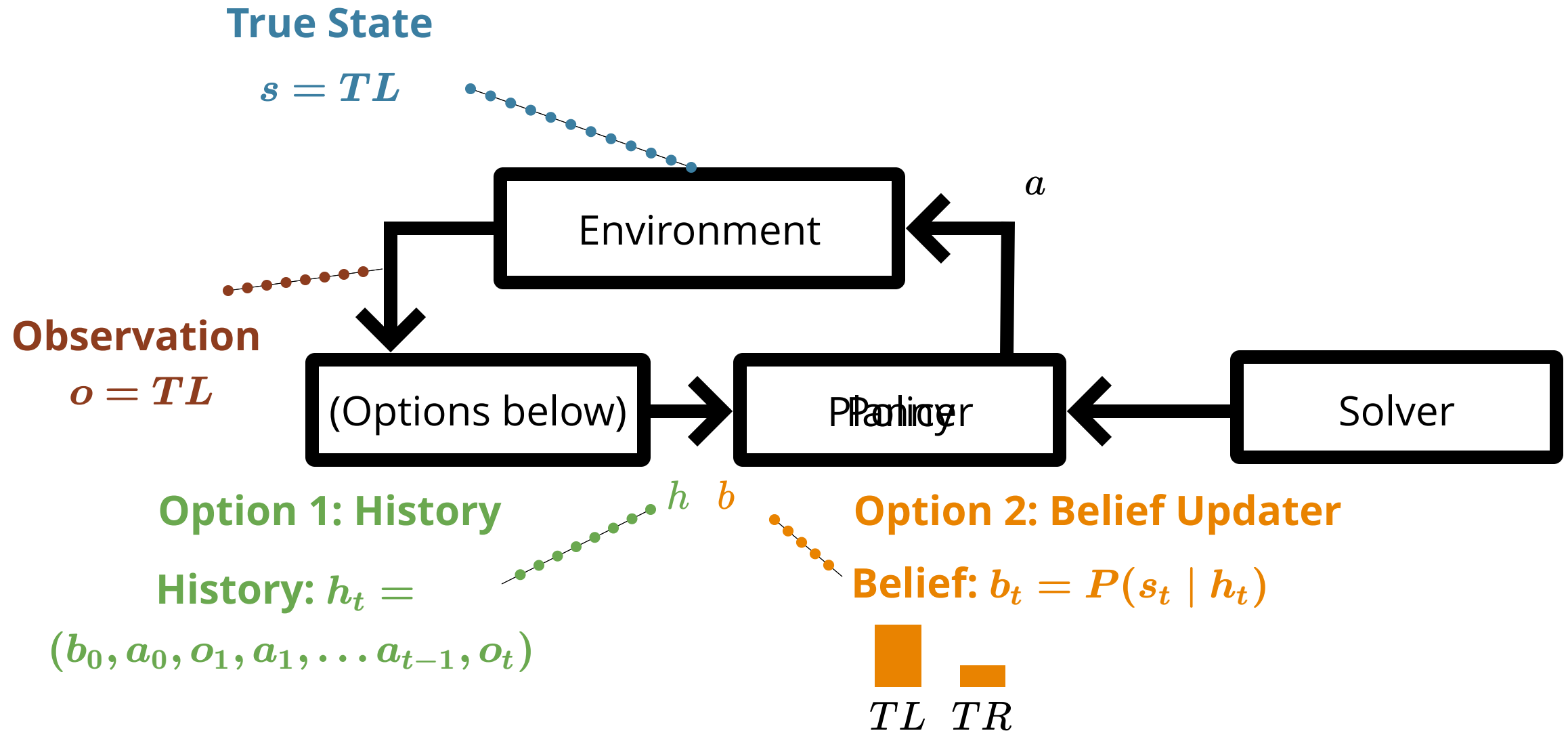
Today!

## Formulation Approximations

(solve a slightly different problem)

Last Time

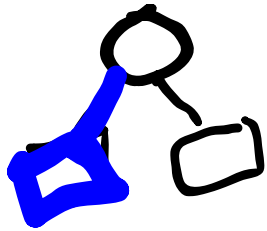
# POMDP Sense-Plan-Act Loop



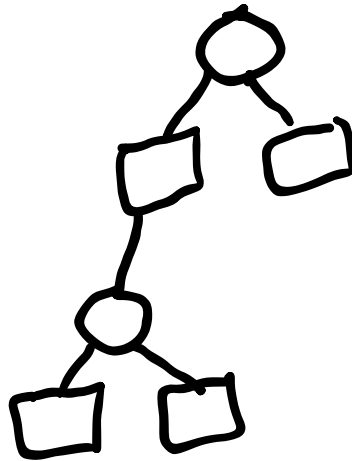
# Belief-Space Tree Search: AEMS

# Monte Carlo Tree Search (MCTS/UCT)

Search



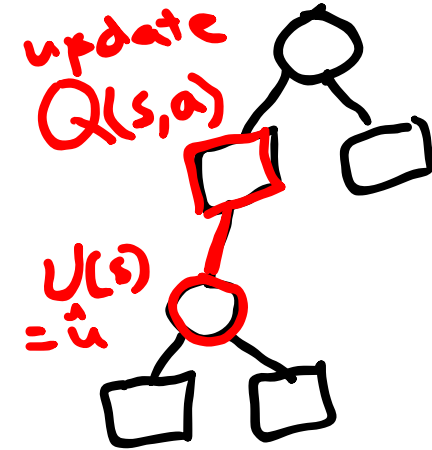
Expansion



Rollout



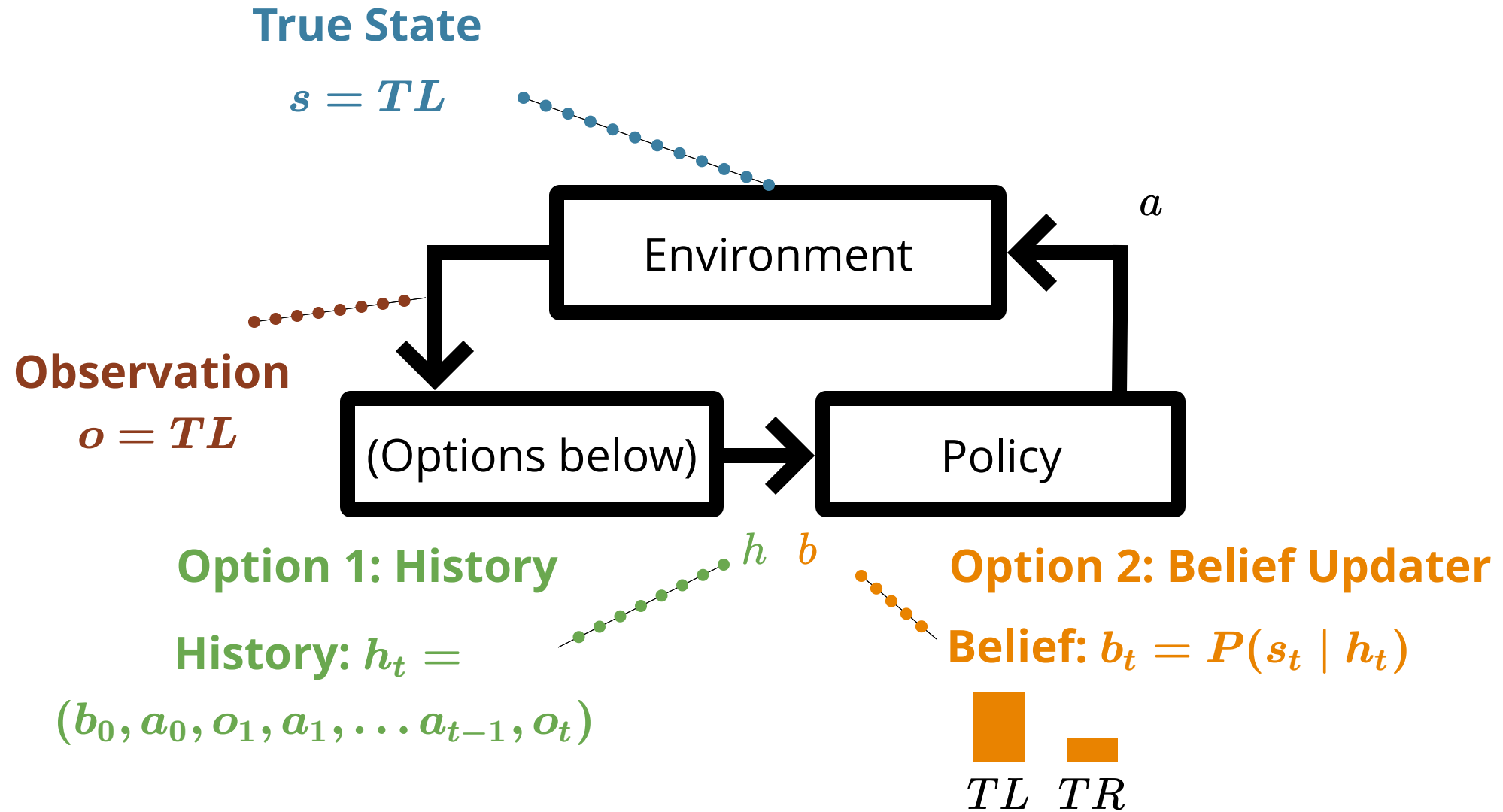
Backup



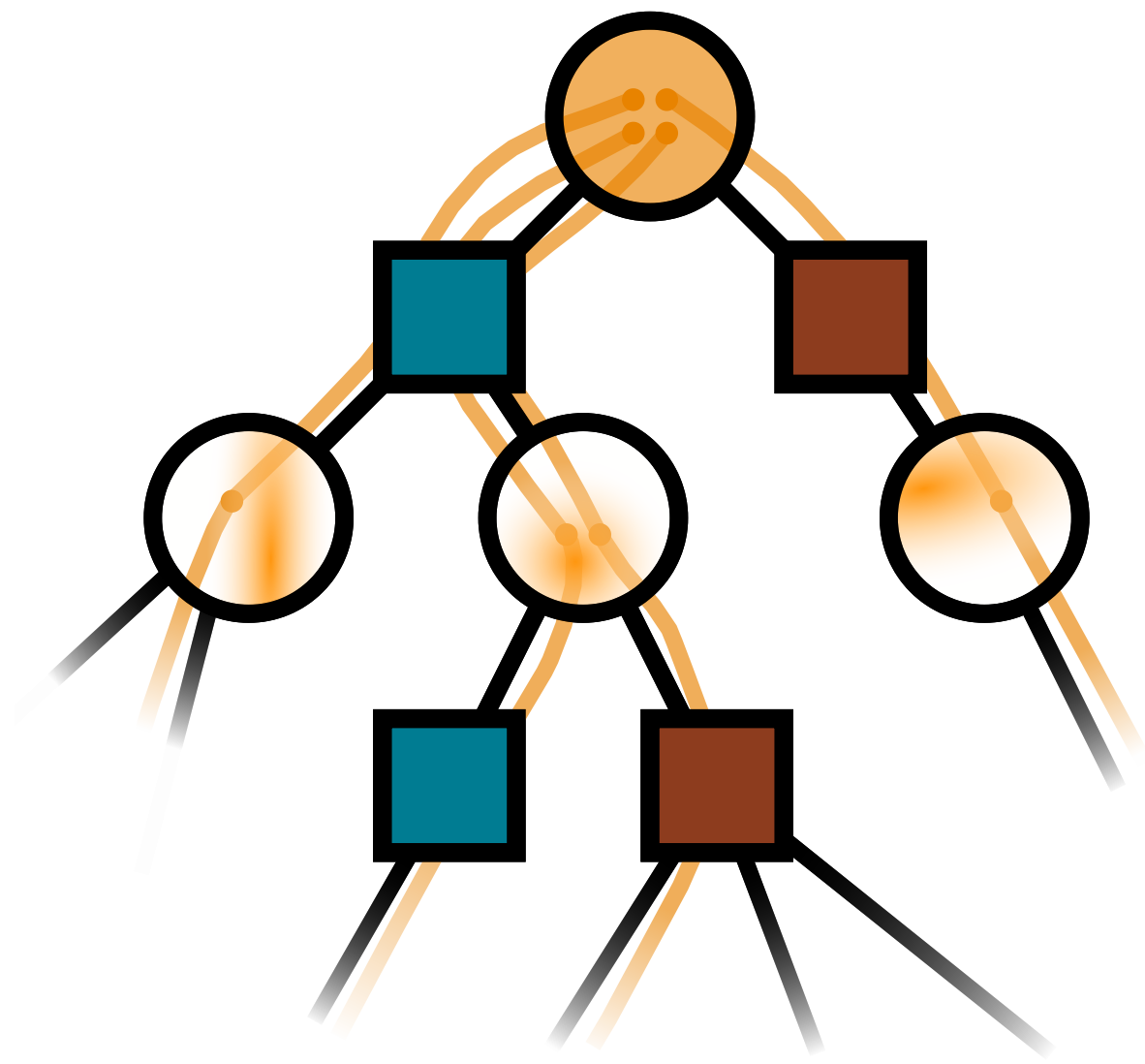
$$Q(s, a) + c \sqrt{\frac{\log N(s)}{N(s, a)}}$$

low  $N(s, a)/N(s)$  = high bonus  
start with  $c = 2(\bar{V} - \underline{V})$

# How should we adapt MCTS for POMDPs?

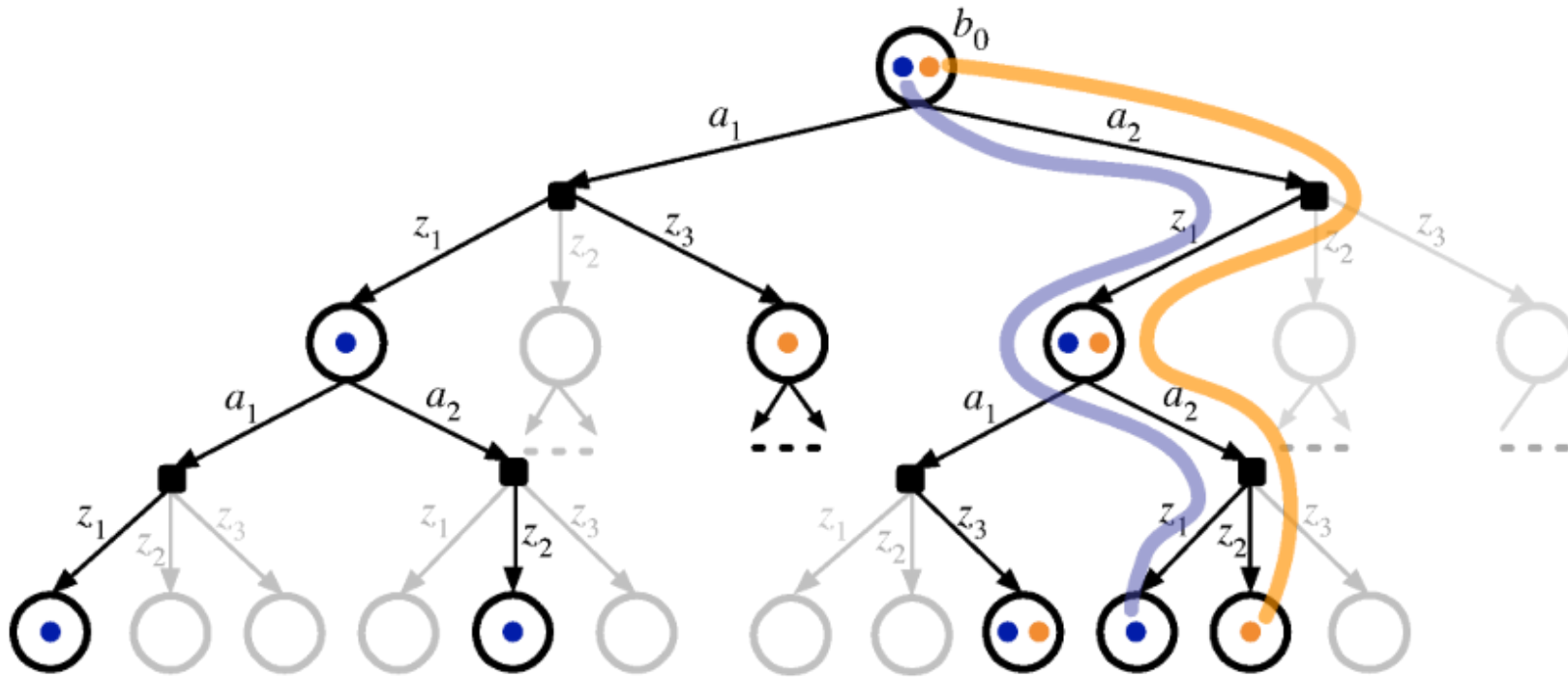


# MCTS on Histories





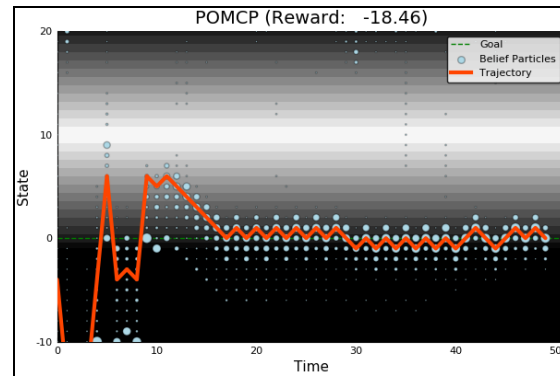
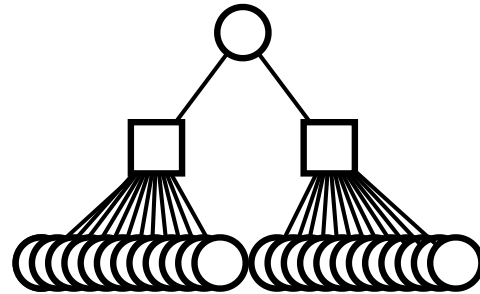
# DESPOT



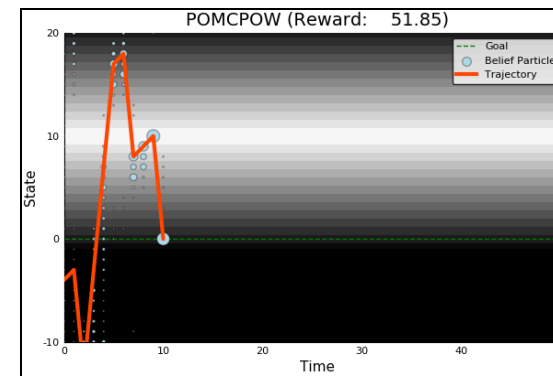
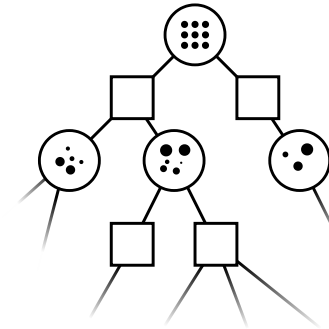
- Determinized Scenarios
- Guided by Lower and Upper Bounds

# Continuous Observation Spaces

## POMCP



## POMCPOW

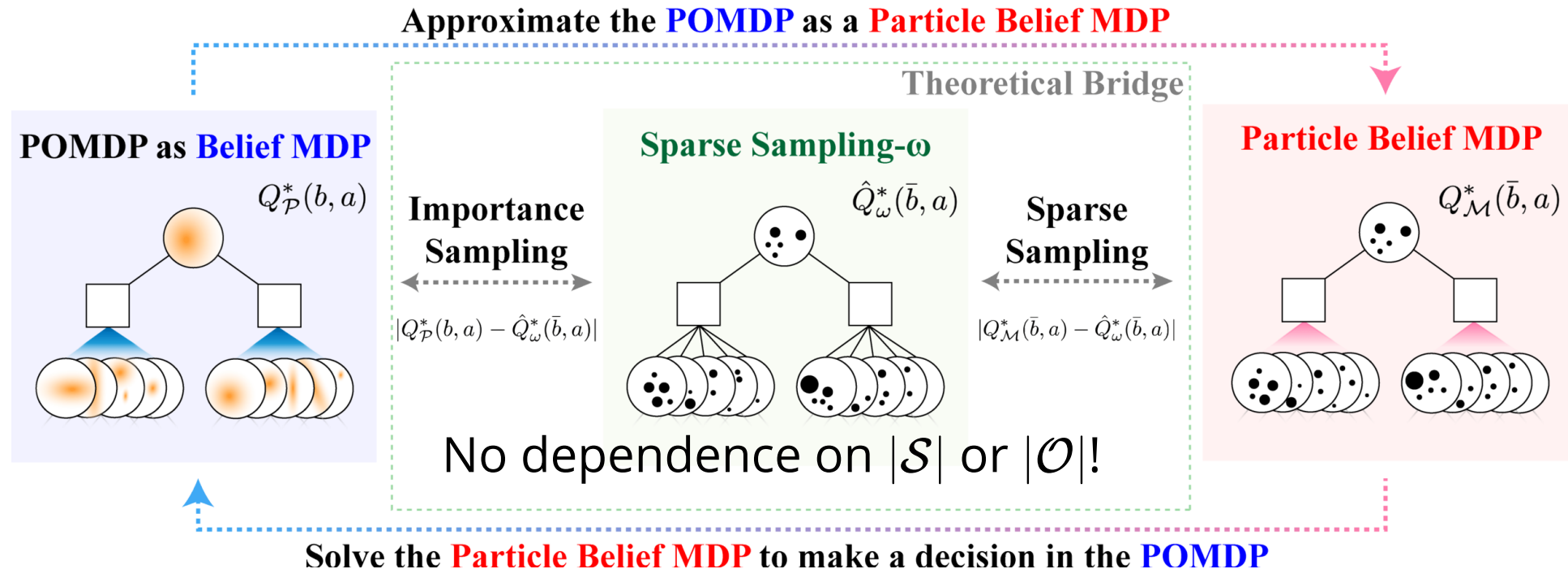


# PF Approximation Accuracy

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

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

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



# DESPOT- $\alpha$

# Continuous Action Spaces

## BOMCP

### Bayesian Optimized Action Branching

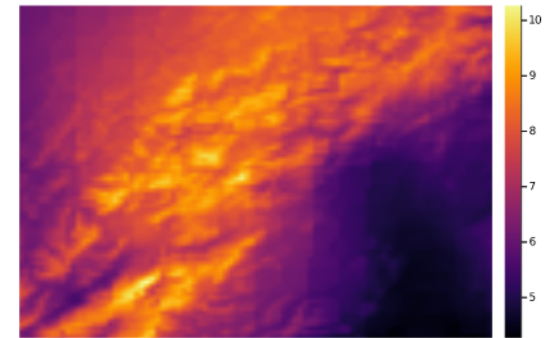
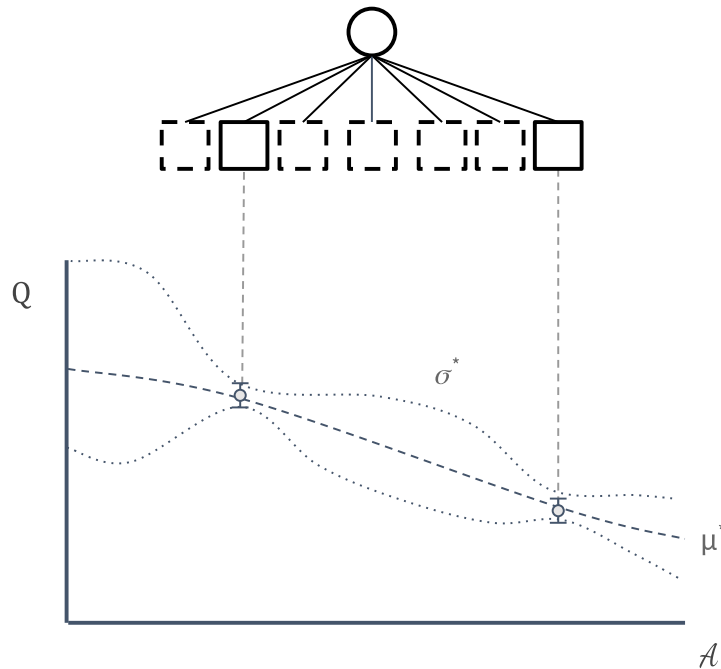
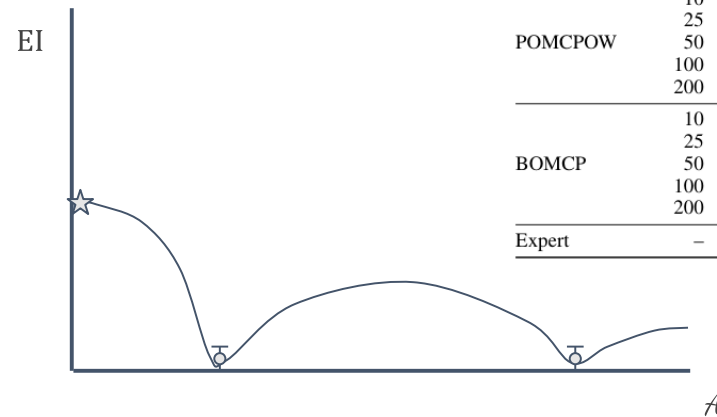


Figure 2: Wind Map. Figure shows wind map for Altamont Pass, CA at 100m altitude. The colors represent the average annual wind speed in m/s.

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



# Voronoi Progressive Widening

