Pareto Monte Carlo Tree Search for Multi-Objective Informative Planning

Weizhe Chen, Lantao Liu
Indiana University Bloomington

I. Summary

- Single-objective informative planning → multi-objective informative planning (e.g. extreme value seeking).
- Pareto MCTS seeks multiple features simultaneously.
- Bounds on the convergence of the tree search process.

II. Background

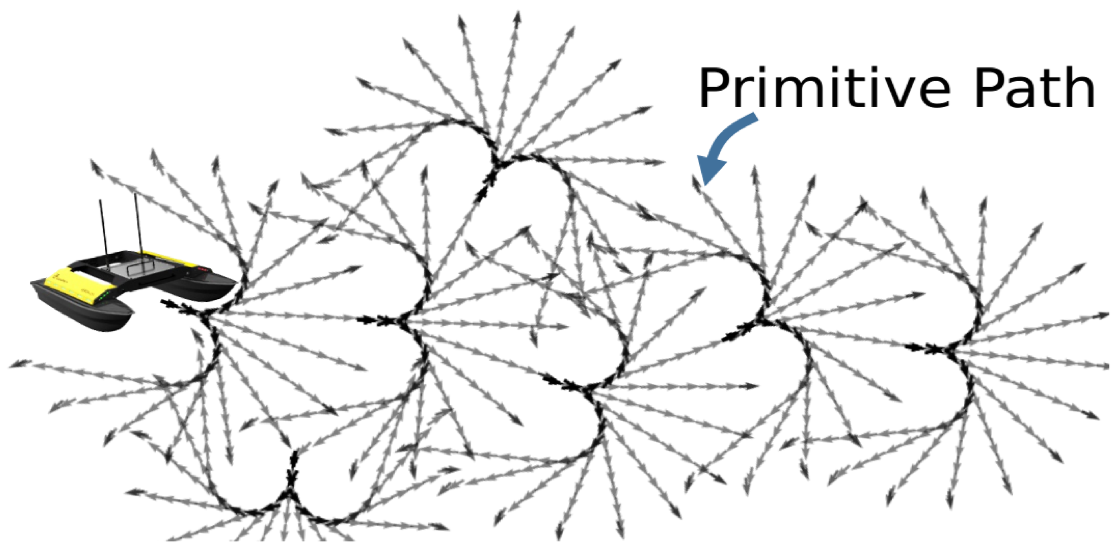
Problem Formulation

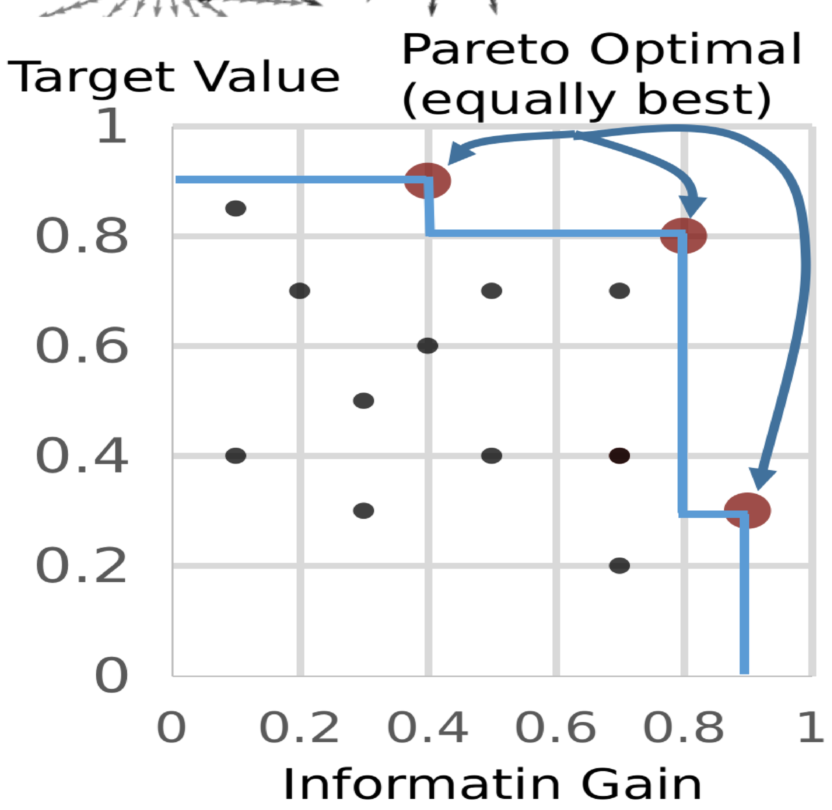
$$\boldsymbol{a}^* = \arg\max_{\boldsymbol{a}\in\mathcal{A}} \{I(\boldsymbol{a}), \underbrace{F_1(\boldsymbol{a}), \ldots, F_{D-1}(\boldsymbol{a})}_{\text{risk, cost, variogram} \ldots}\}$$

s.t. $C_{\boldsymbol{a}} \le B$

- $I(\boldsymbol{a})$: Information gain.
- $F_{\cdot}(\boldsymbol{a})$: Domain-specific goals.

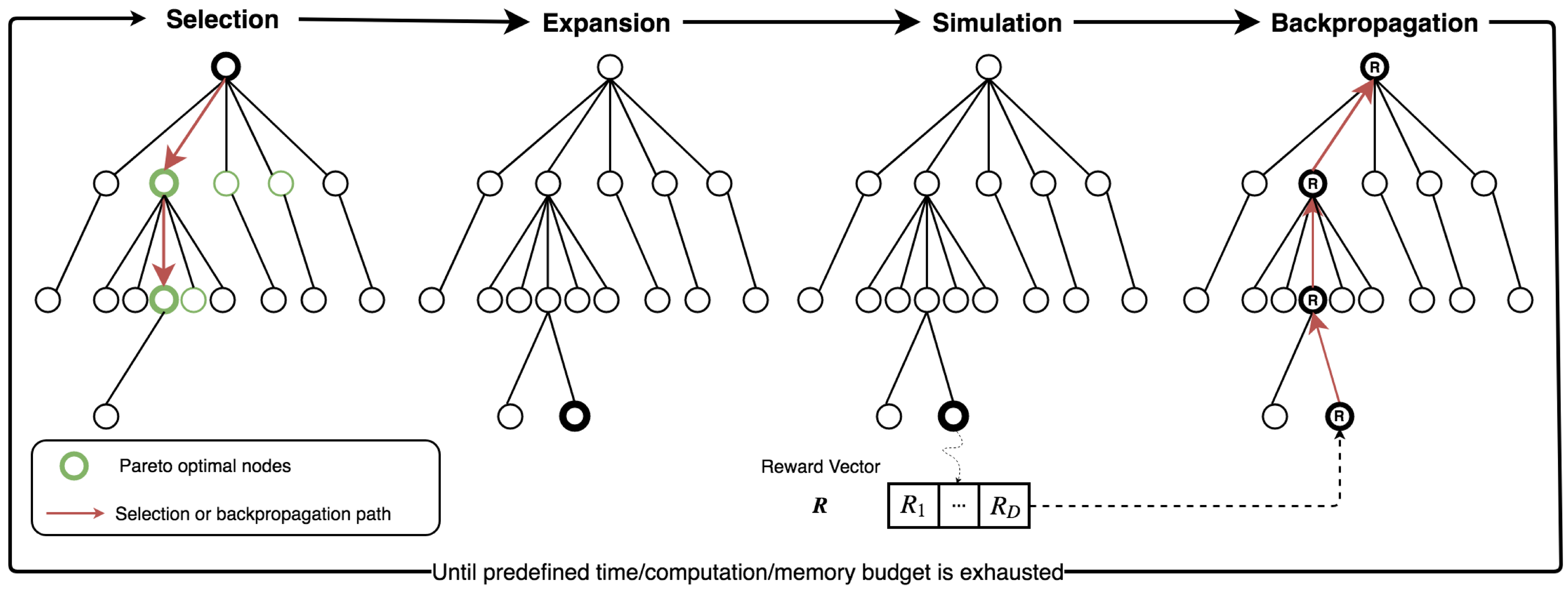
Pareto Optimality





III. Method

Pareto Monte Carlo Tree Search



- Challenge: rewards are vectors → which node is more promising?
  Solution: build Pareto optimal set → randomly expand a node from it.
- Challenge: balance exploitationand exploration in selection step.
  Solution: ·node selection → multi-objective multi-armed bandit.
  ·propose Pareto upper confidence bound (Pareto UCB)

Optimality Analysis of Node Selection Policy

- #sub-optimal nodes selection $\mathbb{E}[T_k(n)]$ is logarithmically bounded:

$$\mathbb{E}[T_k(n)] \le \frac{8\ln n + 2\ln D}{(1-\xi)^2(\min_{k,d}\Delta_{k,d})^2} + N_0(\xi) + 1 + \frac{\pi^2}{3}.$$

- The probability of returning "bad" decision converges to 0:

$$\mathbb{P}(I_t \notin \mathcal{P}^*) \le C t^{-\frac{\rho}{2}\left(\frac{\min_{k,d}\Delta_{k,d}}{36}\right)^2}.$$

IV. Experiments

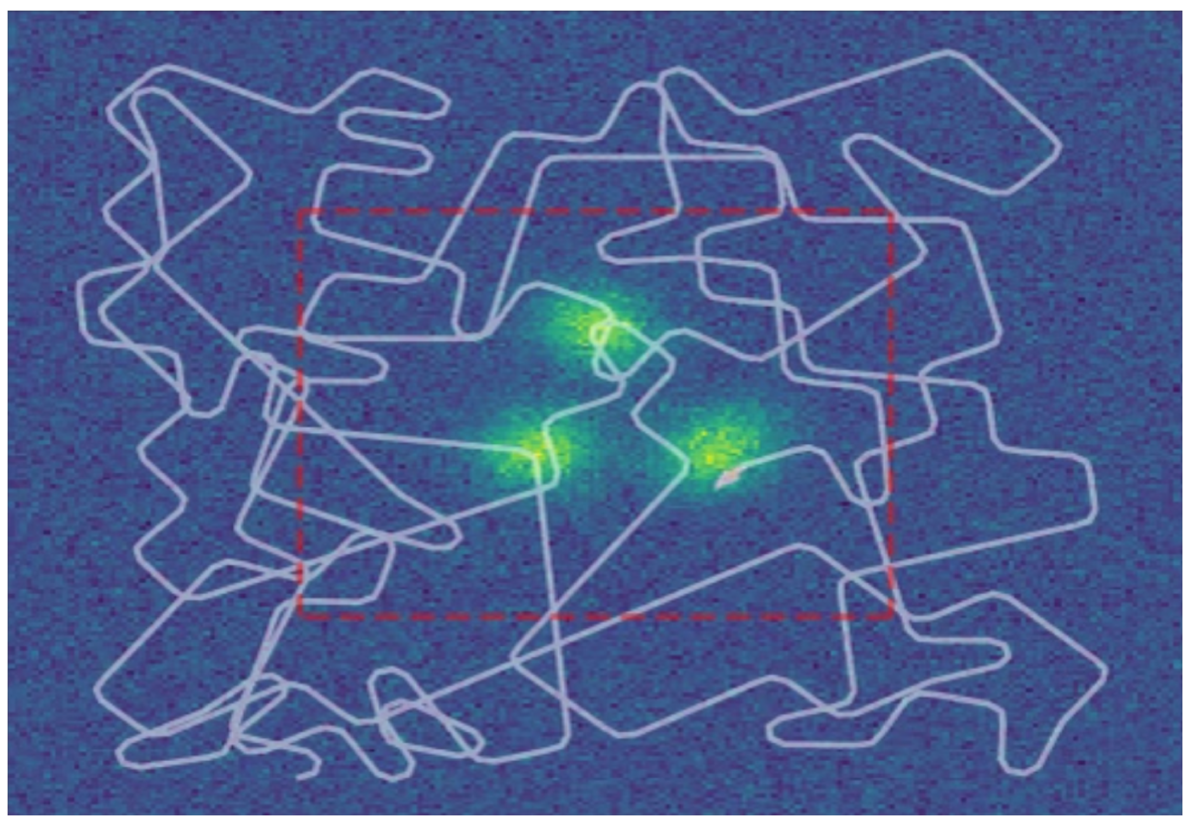


Figure 1: Information MCTS

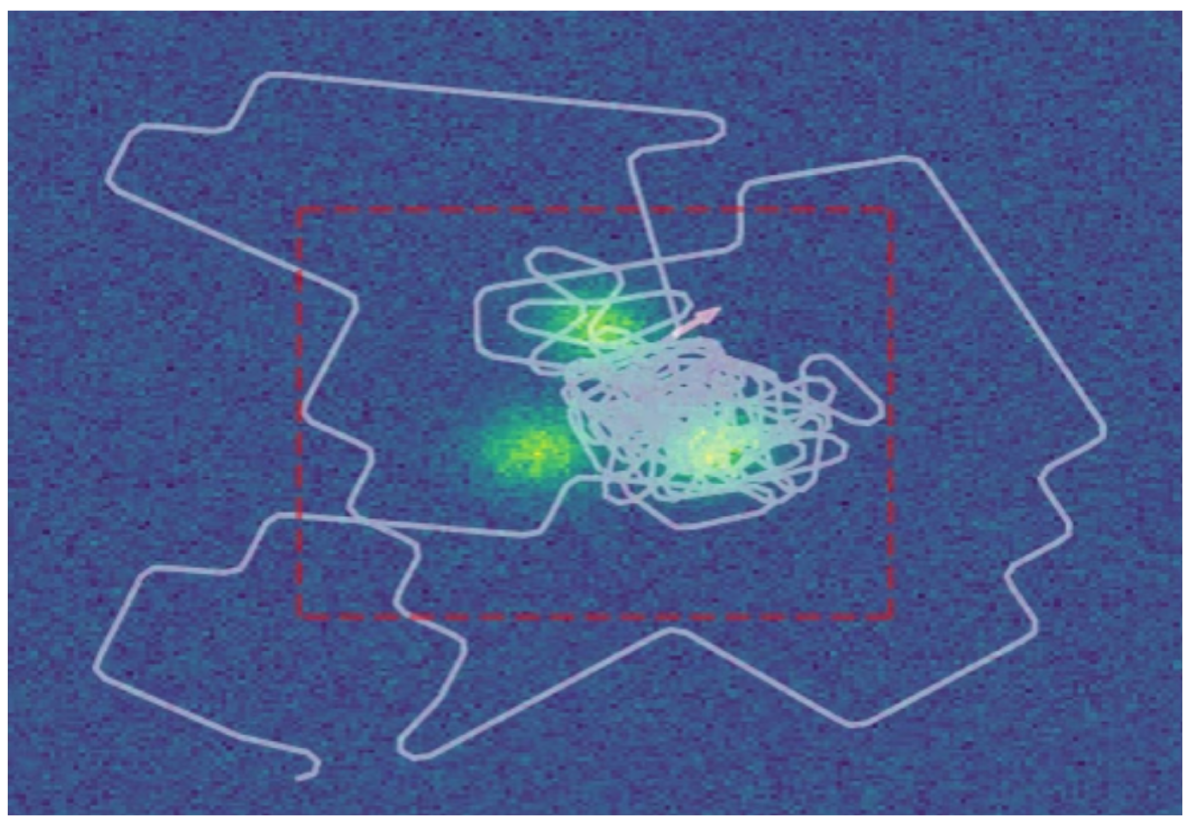


Figure 2: UCB MCTS

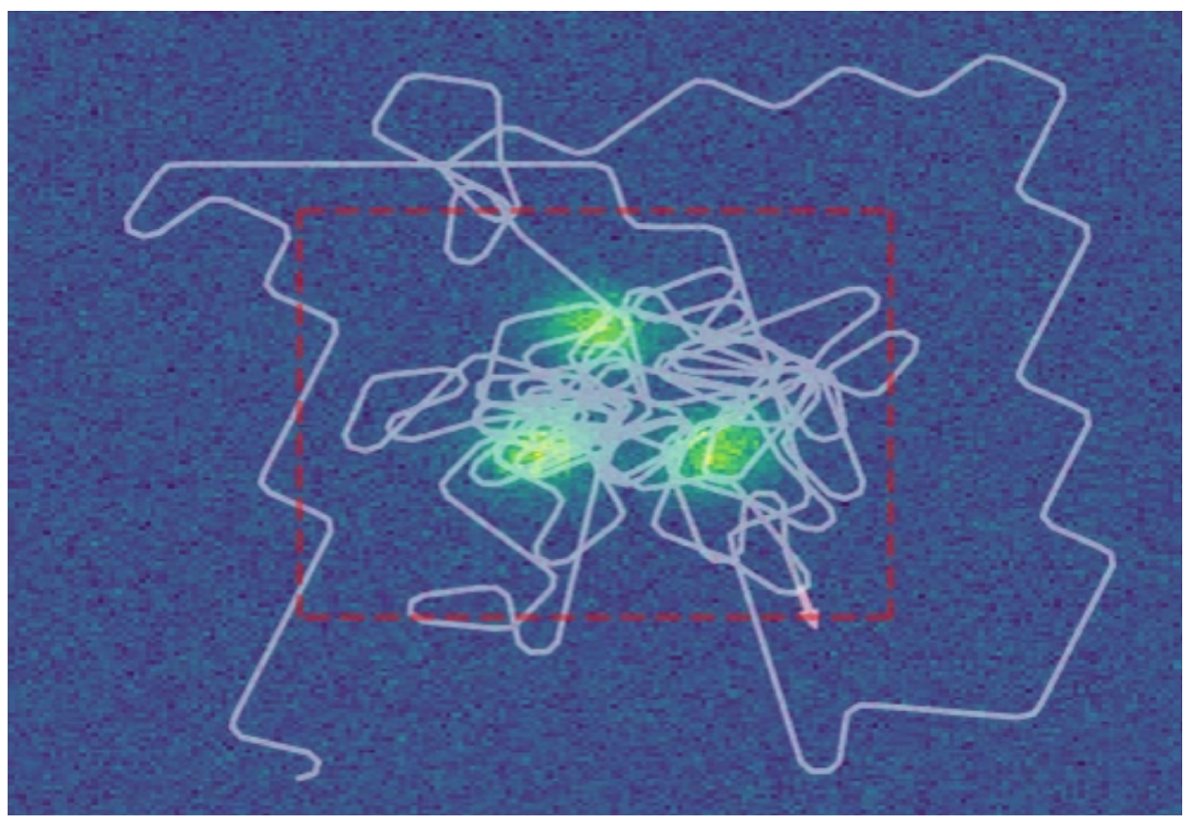


Figure 3: Pareto MCTS

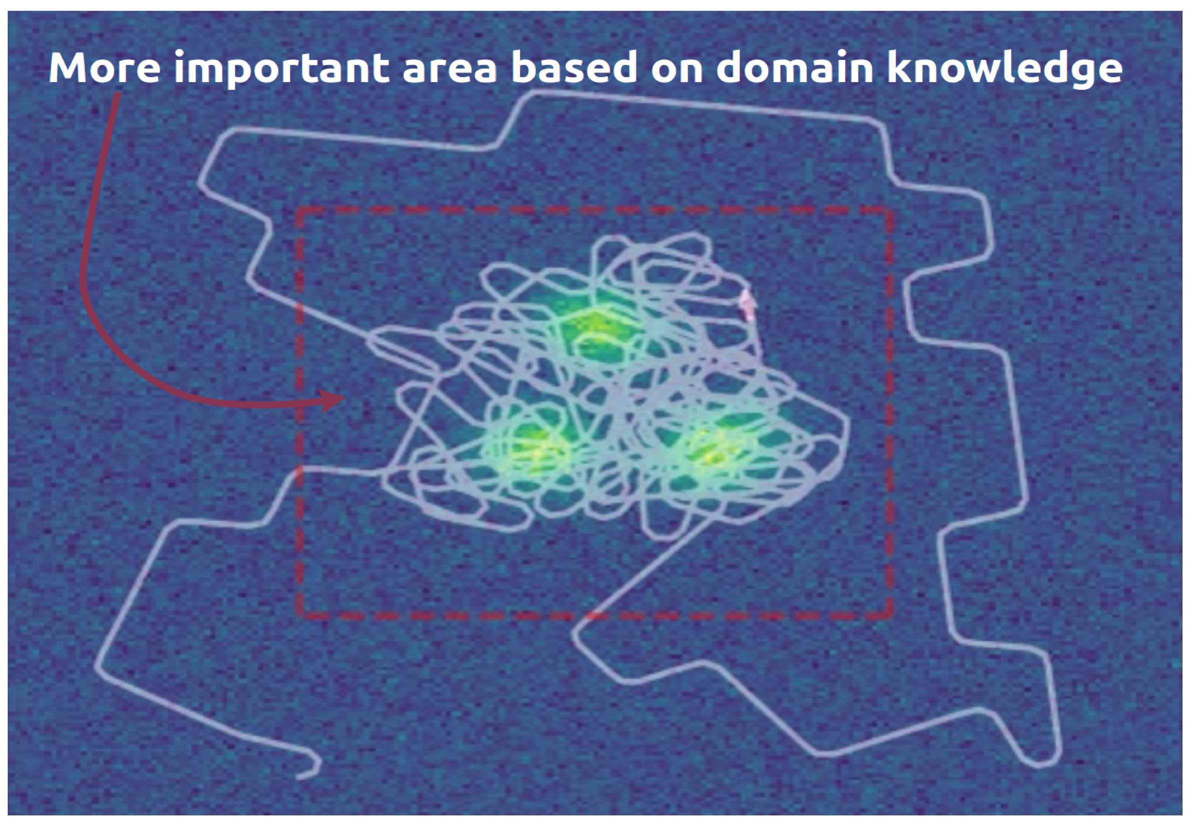


Figure 4: With Expert Knowledge

V. Remarks

- Multi-objective informative planning enables robots to exhibit more behaviors other than exploration.
- Expert knowledge could be included by user-defined objective functions or biased node selection.