## Planning in an Abstract Space



Subgoals → Abstract MDP → Subgoal Values $v_{g^\star}(s)$

Option Policies → Local Models $r(s,g), \Gamma(s,g)$ $\tilde{r}(g,g'), \tilde{\Gamma}(g,g')$
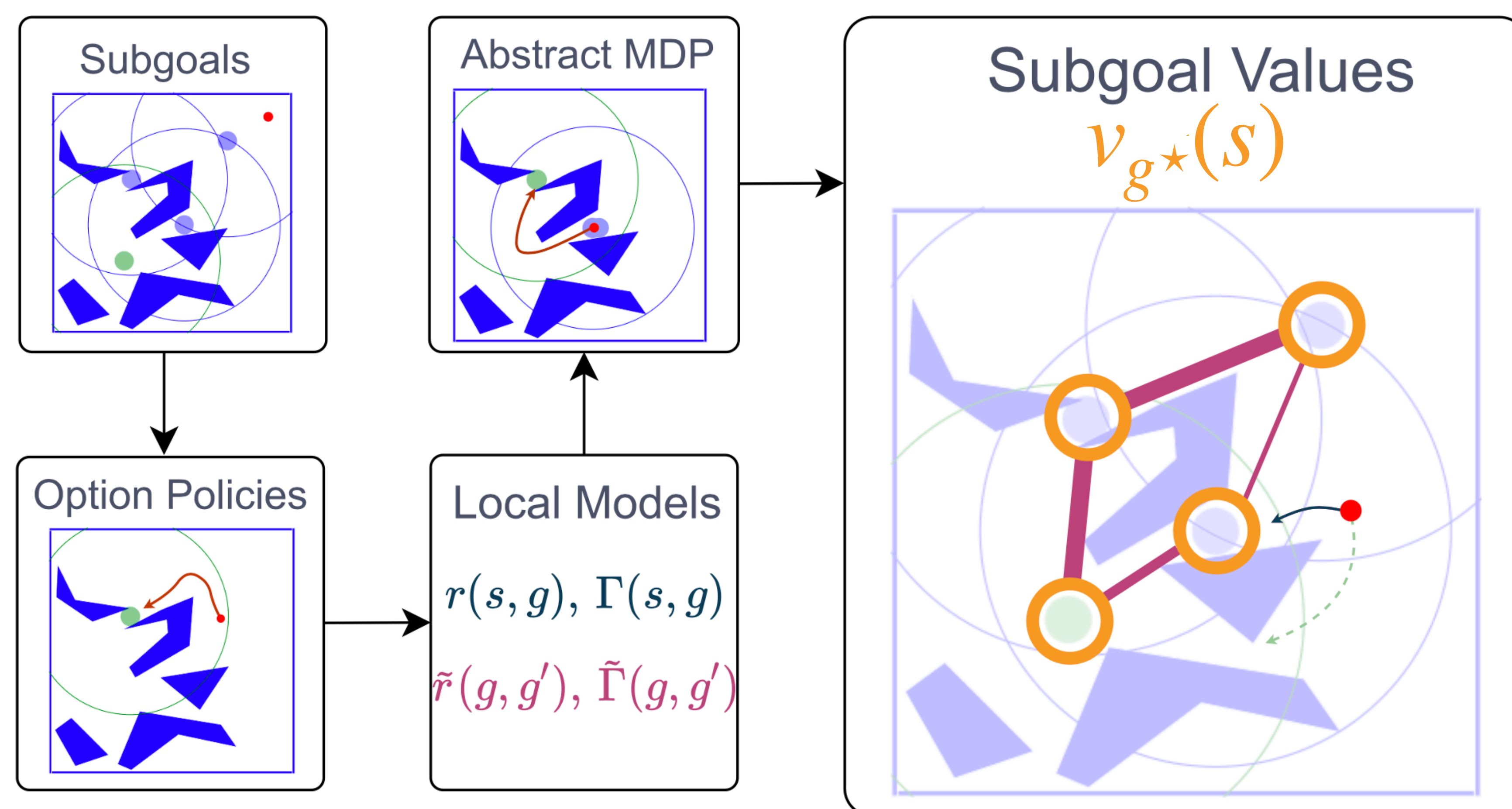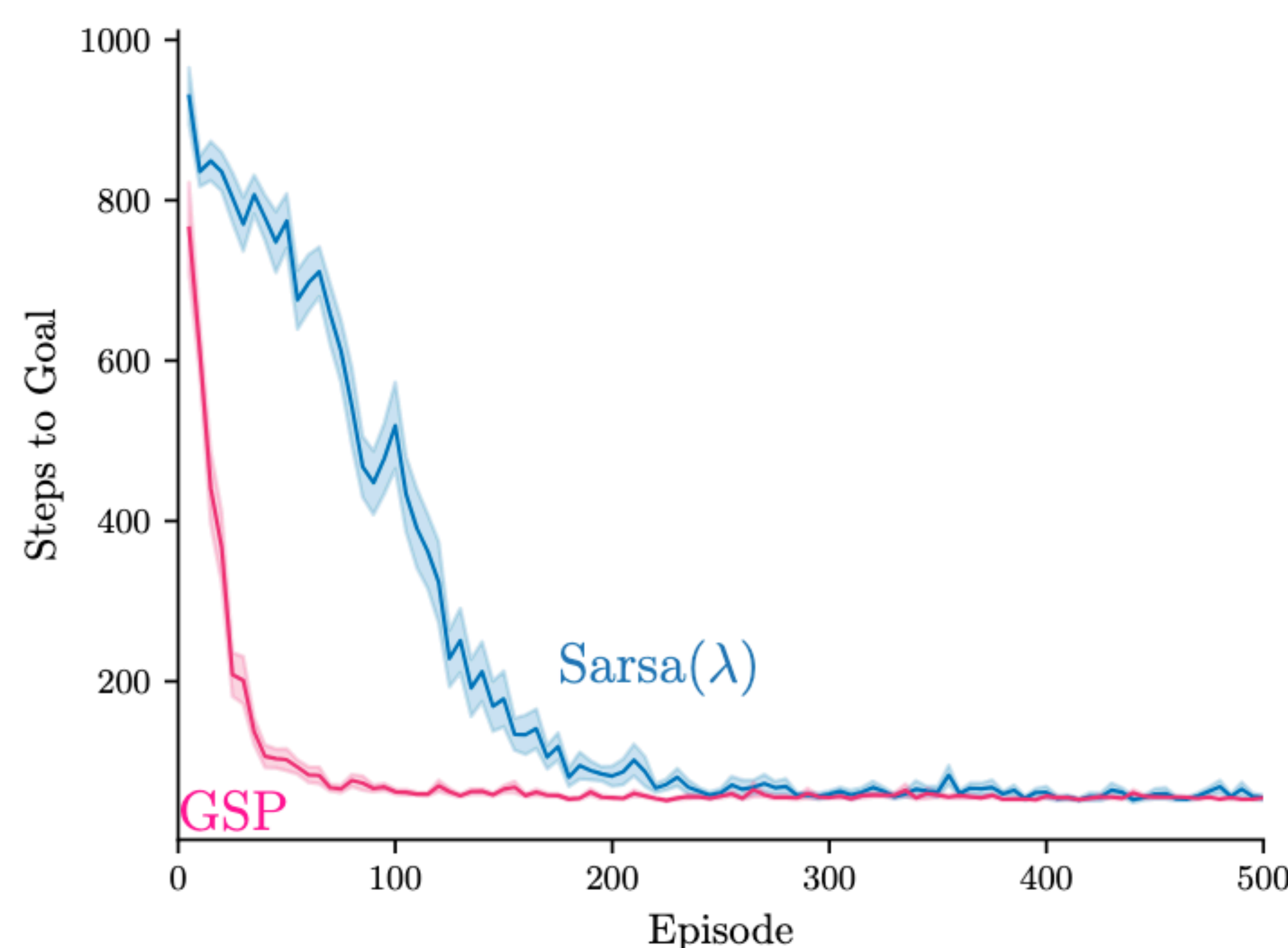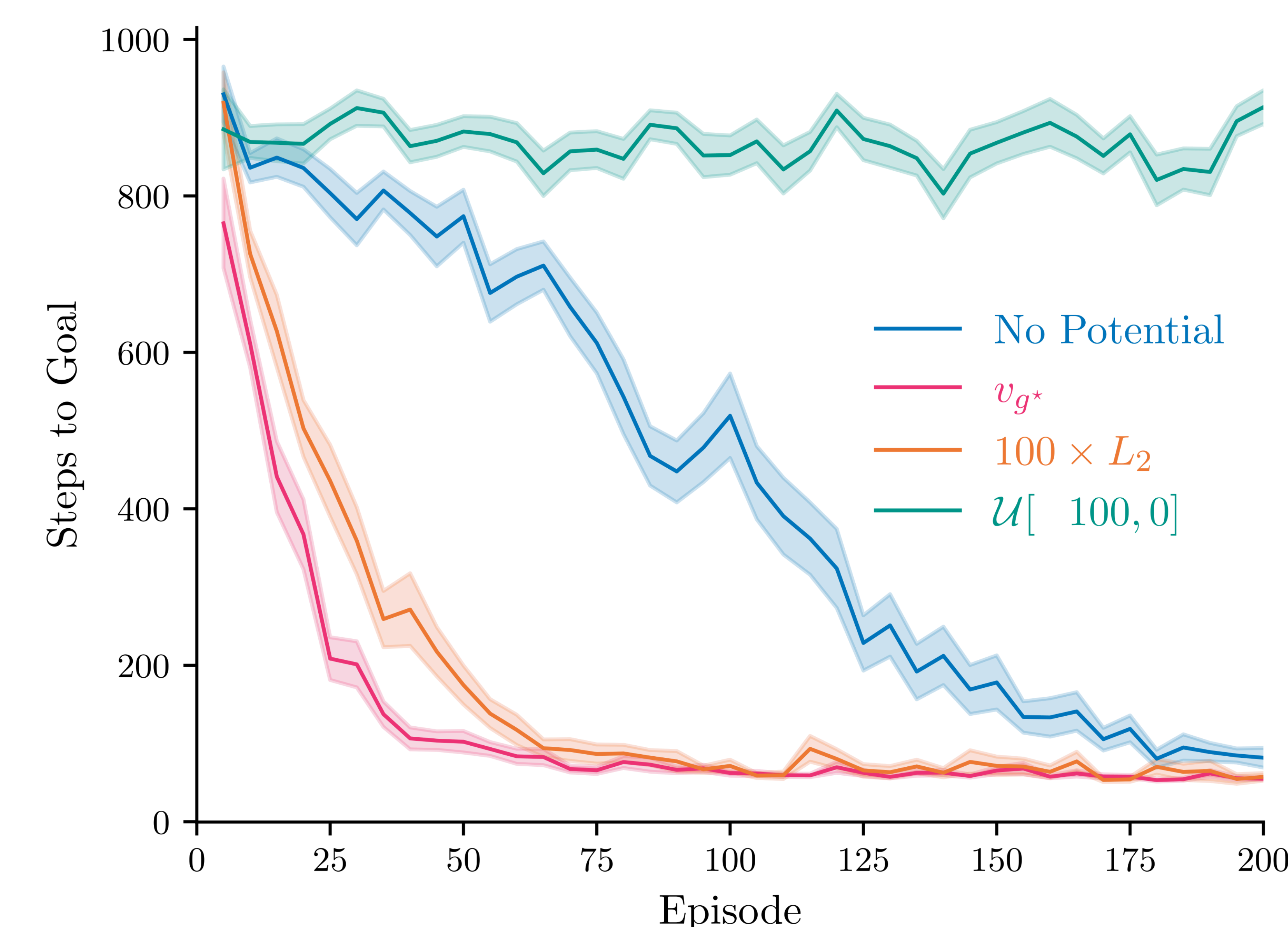
- Planning in large Markov Decision Processes (MDPs) is expensive.
- We construct an Abstract MDP of subgoals and options.
- We learn return and discounted probability models of these options.
- Perform value iteration in the abstract MDP and project values down to the original space for reward shaping.
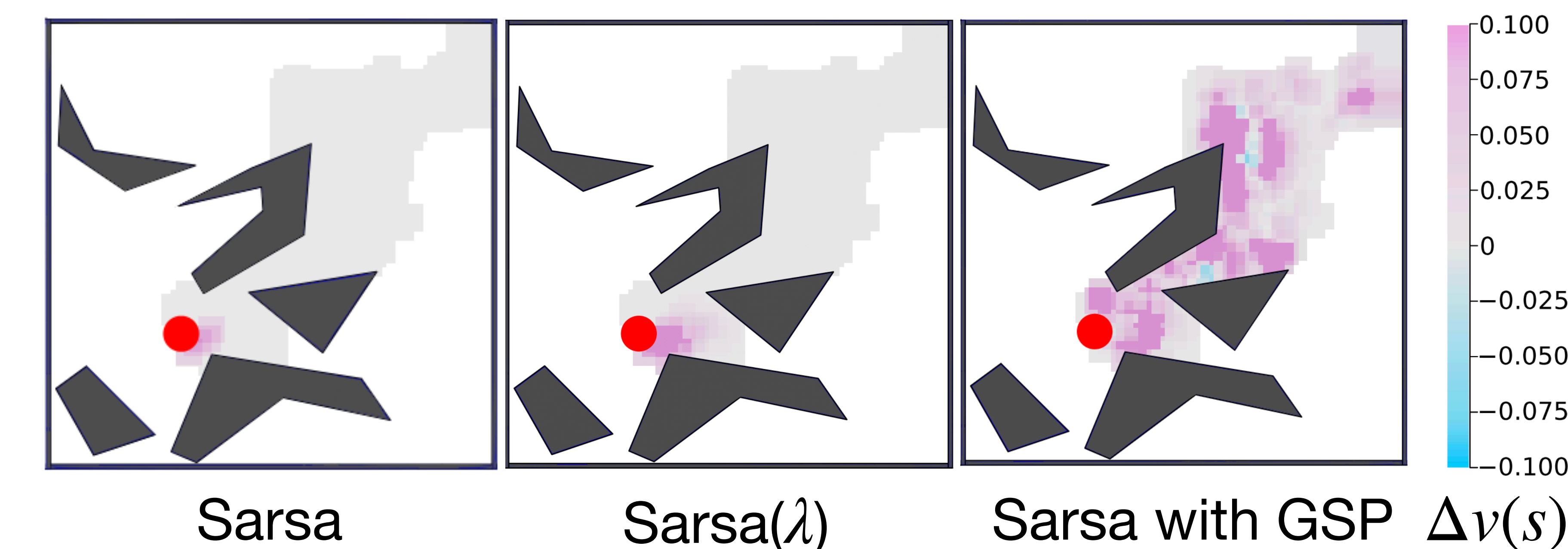
## Goal-Space Planning is Faster



Sarsa($\lambda$)

GSP

## Abstract models can be used for planning by reward shaping.

## Shaping with World Models



No Potential
$v_{g^\star}$
$100 \times L_2$
$\mathcal{U}[\,100, 0]$

## GSP Propagates Value



Sarsa    Sarsa($\lambda$)    Sarsa with GSP    $\Delta v(s)$

## Limitations

- Computational cost of model learning.
- Discovering subgoals for planning.
- Shaping is not as helpful in trajectory-based methods (REINFORCE or PPO)

### Goal-Space Planning with Subgoal Models

Chunlok Lo[=12], Kevin Roice[=12], Parham M. Panahi[=12], Scott M. Jordan[12], Adam White[123], Gabor Mihucz[12], Farzane Aminmansour[12], Martha White[123]

1 Department of Computing Science, University of Alberta
2 Alberta Machine Intelligence Institute
3 Canada CIFAR AI Chair