

Learning and Optimizing Probabilistic Models for Planning under Uncertainty

R. van Bekkum

M. T. J. Spaan

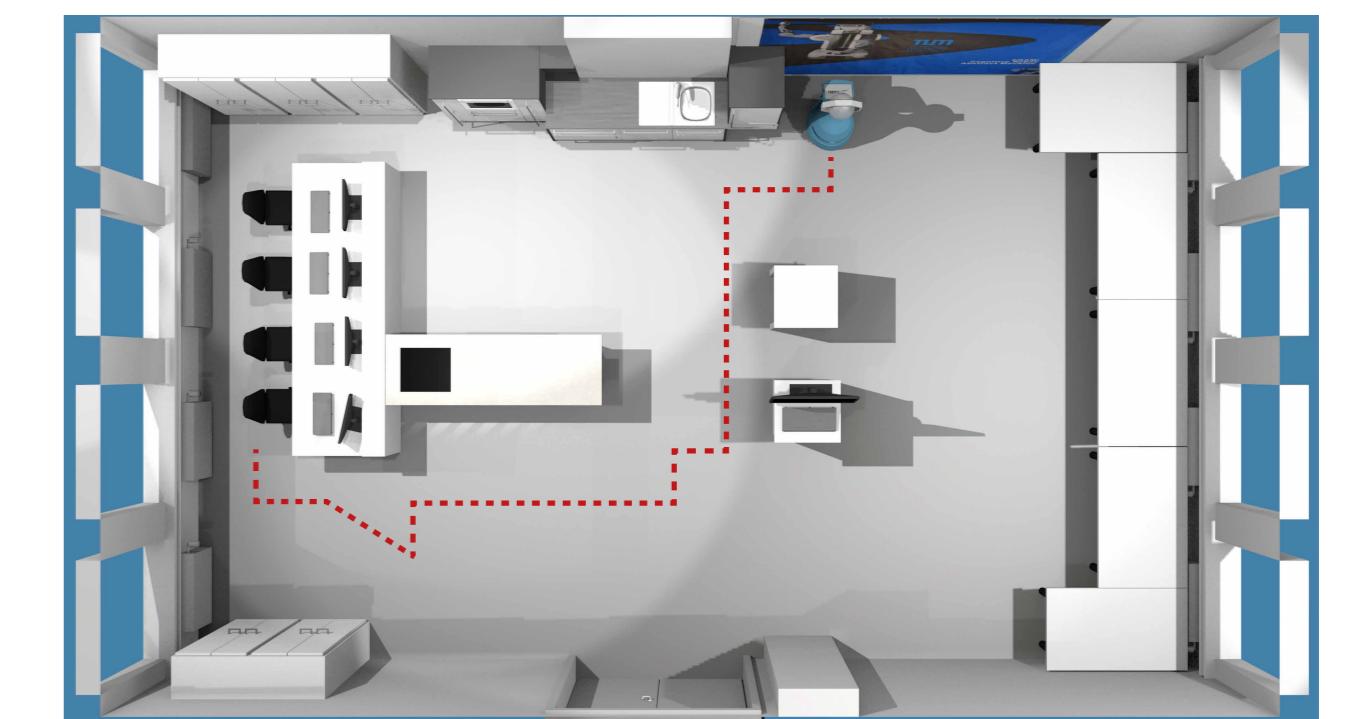
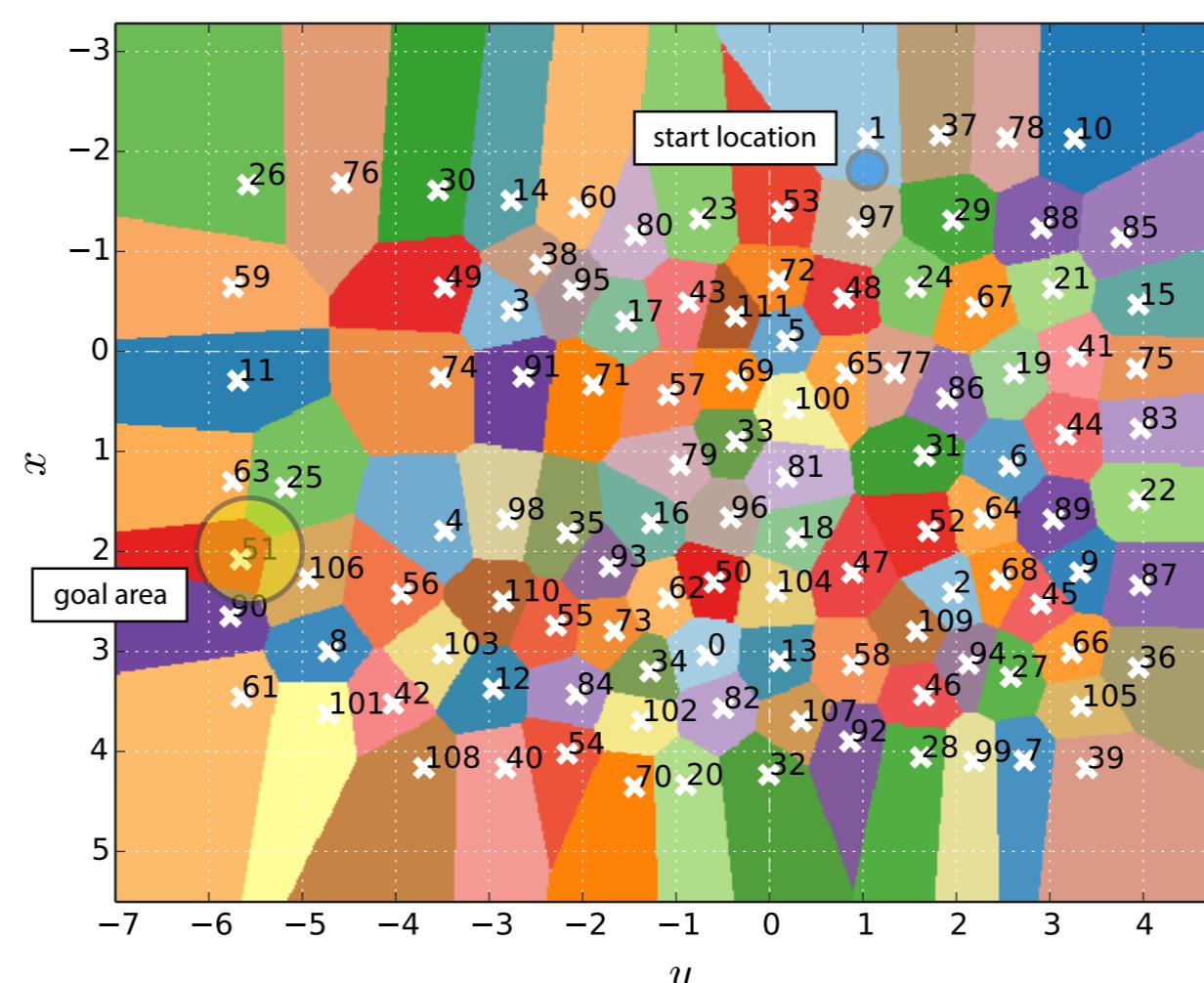
Algorithmics Group, Department of Software Technology

Abstract

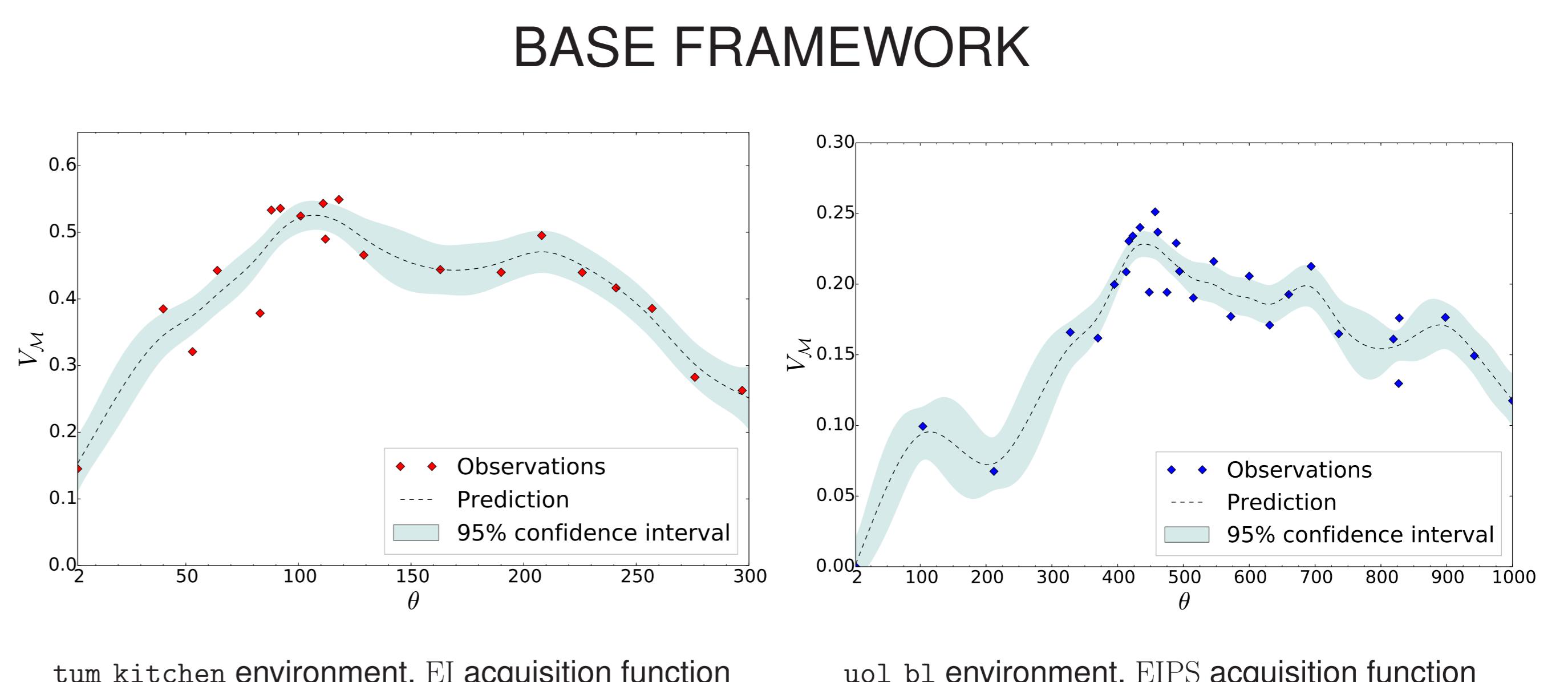
Decision-theoretic planning techniques are increasingly being used to obtain (optimal) plans for domains involving uncertainty, which may be present in the form of the controlling agent's actions, its percepts, or exogenous factors in the domain. These techniques build on detailed probabilistic models of the underlying system, for which Markov Decision Processes (MDPs) have become the *de facto* standard formalism. However, handcrafting these probabilistic models is usually a daunting and error-prone task, requiring expert knowledge on the domain under consideration. Therefore, it is desirable to automate the process of obtaining these models by means of learning algorithms presented with a set of execution traces from the system. Although some work has already been done on crafting such learning algorithms, the state of the art lacks an automated method of configuring their hyperparameters, so to maximize the performance yielded from executing the derived plans. In this work we present a solution that employs the *Bayesian Optimization* (BO) framework to learn MDPs autonomously from a set of execution traces, optimizing the expected value and performance in simulations over a set of tasks the underlying system is expected to perform. The approach has been tested on learning MDPs for mobile robot navigation, motivated by the significant uncertainty accompanying the robots' actions in this domain.

Experimental Setup and Results

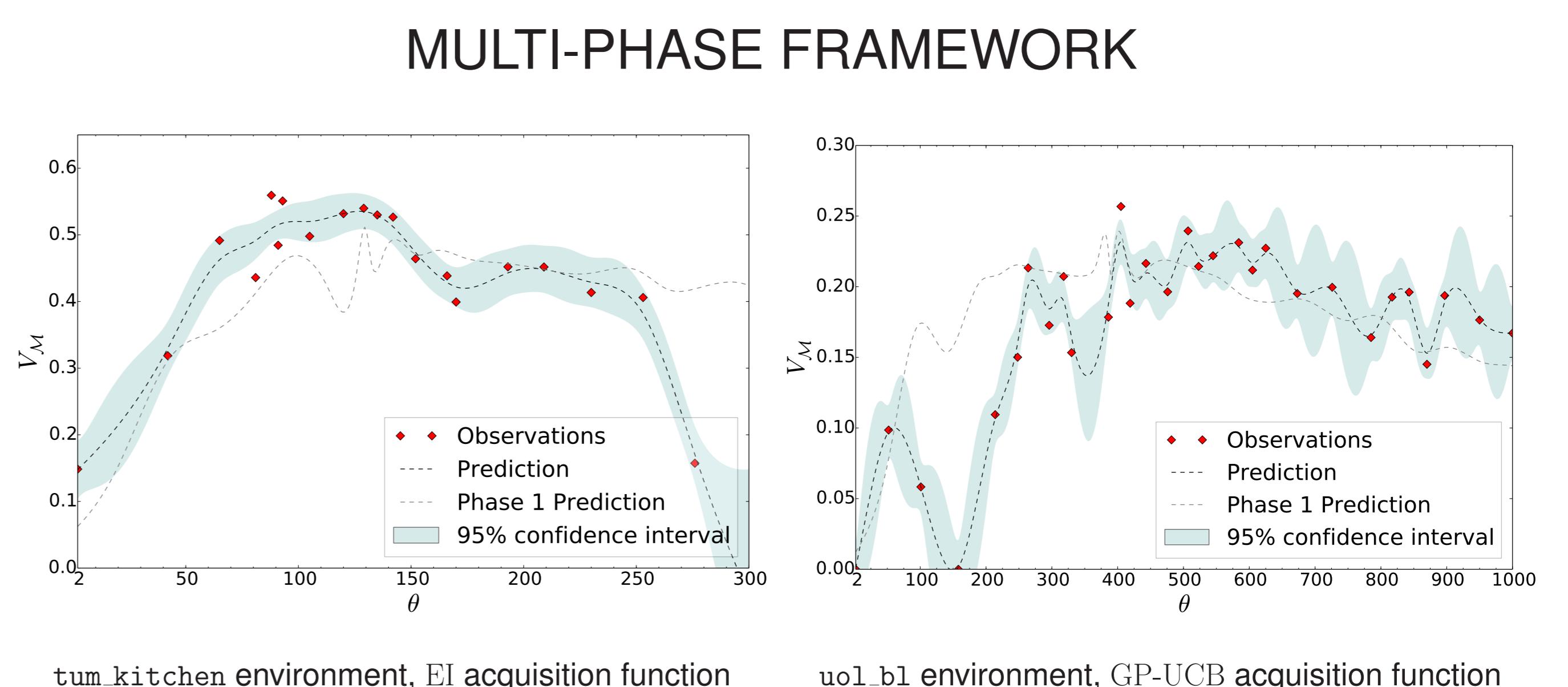
- Dataset of execution traces with odometric readings
- Learning MDPs by k -Means and Maximum Likelihood



- Navigation in simulations of small and large environments
- Extract hyperparameters from resulting GP posteriors in BO



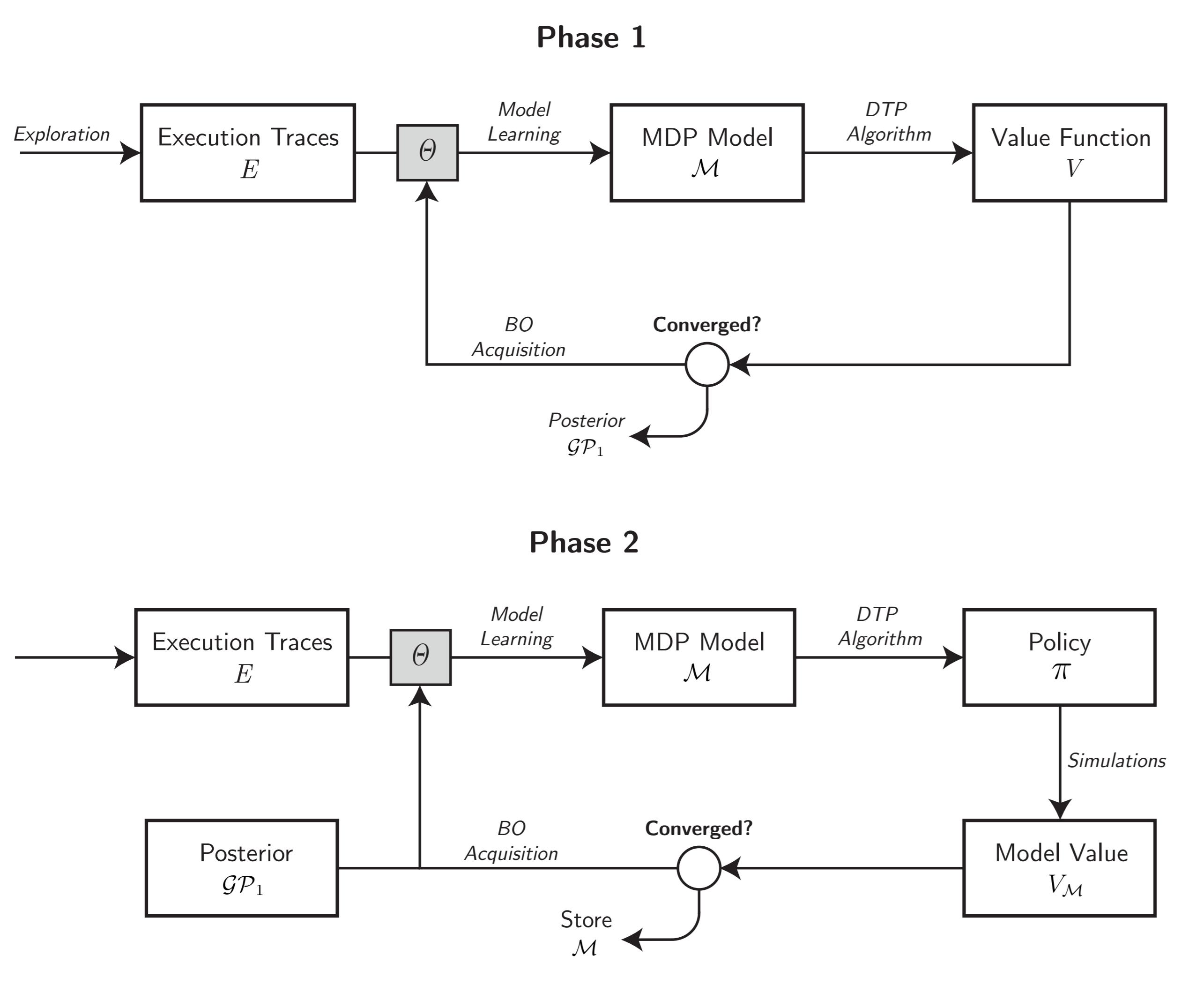
BASE FRAMEWORK



MULTI-PHASE FRAMEWORK

Methodology

- BO for the parameters of MDP model learning algorithms
- Performance assessed through expensive simulations
- Multi-phase for speeding up BO with informative prior



Conclusions

- Evaluating all parameter settings of model learning algorithms can be a time-expensive endeavour, potentially more effectively approached through BO.
- The multi-phase approach may further reduce time-expenses by optimization with cost-cheap performance estimates on a more abstract level.