

A TEST ENVIRONMENTS

A.1 ENVIRONMENTS A-D $15m \times 15m$

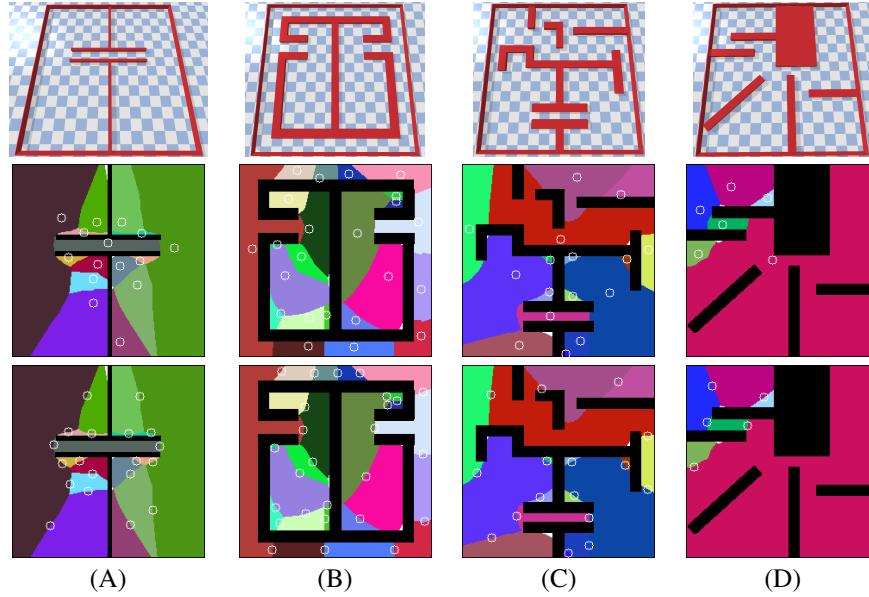


Figure 4: Test environments of the size $15m \times 15m$ with the identified abstract states. These images show 2D projections of high-dimensional region-based Voronoi diagrams. Each colored partition represents an abstract state. Top: The white circles represent centroids of the predicted critical regions used to synthesize centroid options. Bottom: The white circles represent the interface regions for each pair of abstract states used to synthesize interface options.

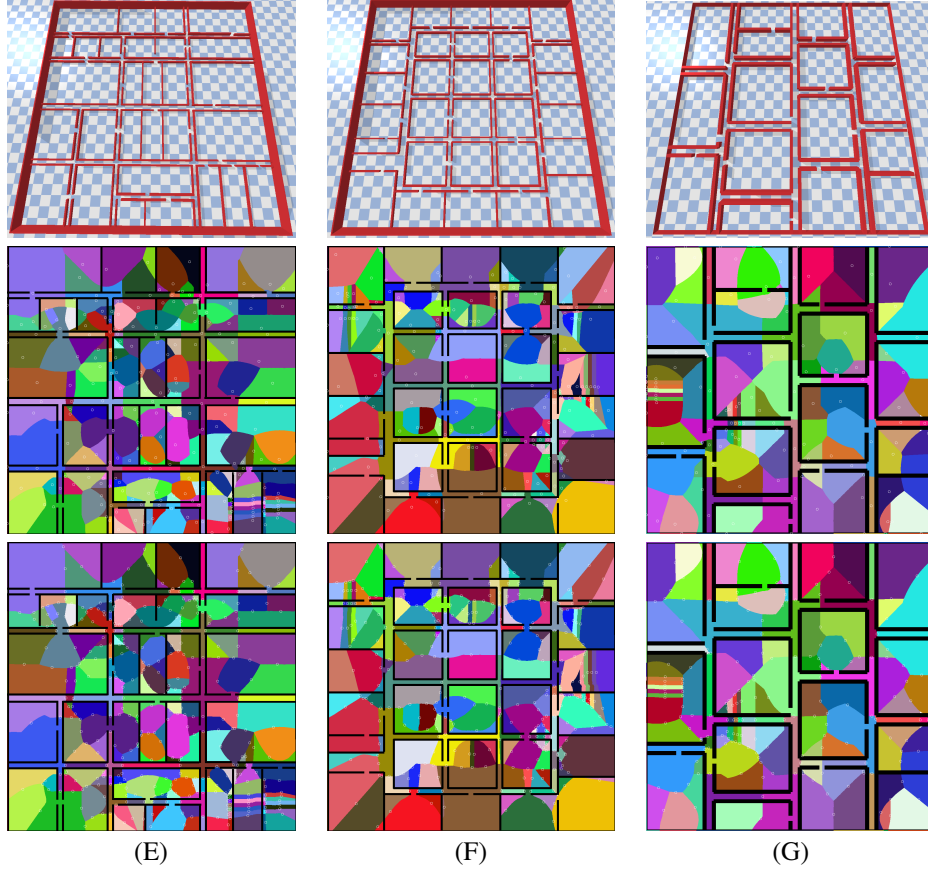
A.2 ENVIRONMENTS E-G $75m \times 75m$ 

Figure 5: Test environments of the size $75m \times 75m$ with the identified abstract states. These images show 2D projections of high-dimensional region-based Voronoi diagrams. Each colored partition represents an abstract state. Top: The white circles represent centroids of the predicted critical regions used to synthesize centroid options. Bottom: The white circles represent the interface regions for each pair of abstract states used to synthesize interface options.

B ENV D RESULTS

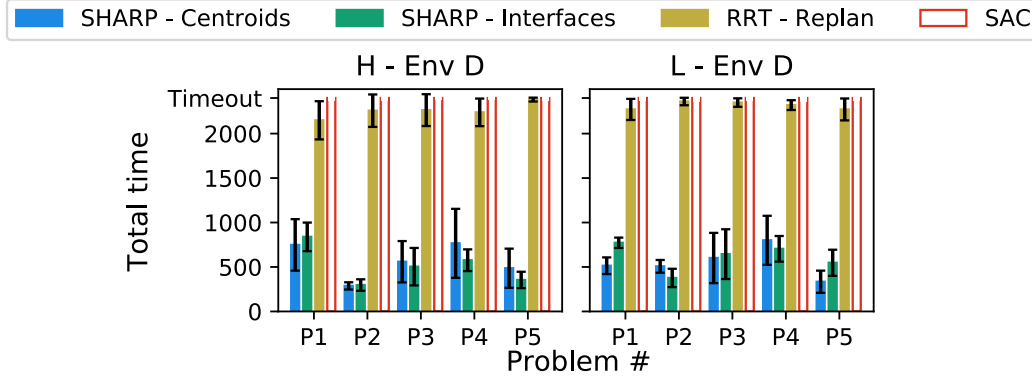


Figure 6: The figure shows the time taken by our approach and baselines to compute path plans in the test environment D. The x-axis shows the problem instance and the y-axis shows the time in seconds. The reported time for our approach includes time to predict critical regions, construct abstractions, and learn policies for all the options. Each subsequent problem instance uses trained policies for options from the previous problems if there exists one. Timeout was set to 2400 seconds. The numbers are averaged over 5 independent trails. The transparent bars for SAC show that training was stopped as it reached the timeout.

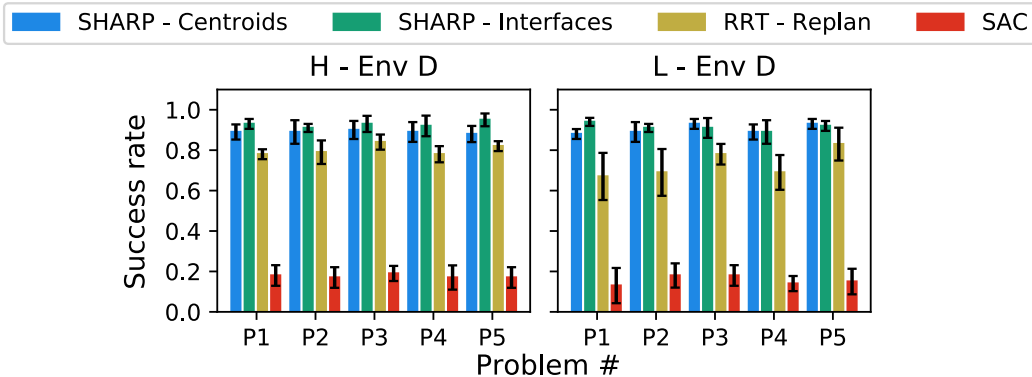


Figure 7: The figure shows the success rate of our approach and baselines in the test environments. The x-axis shows the problem instance and the y-axis shows the fraction of successful executions of the model out of 20 test executions. We used the final policy for our approach and SAC. RRT computed a new plan for each execution with a timeout set to 2400 seconds. The numbers are averaged over 5 independent trails.

C OPTION REUSE RATES

	Env A	Env B	Env C	Env D	Env E	Env F	Env G
Centroid Options	50%	50%	33%	33%	39%	36%	50%
Interface Options	43%	33%	20%	33%	37%	33%	42%

Table 1: Percentage of options that we reuse by our approach across P1-P5.