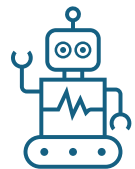


Expert Model



Policy Model

② Expert  
Generated  
Alternatives

① Failed Trajectory

③ Candidate Critical Step  
Selection via PRM Scoring

④ Branch Rollout and  
Outcome Verification

Verified Critical Step  
Collection

⑥ Critical Step  
Optimization (CSO)

⑤ Verified Critical Step  
Preference Data

