

CONSTRUCT ENVIRONMENT

ACTION **OBS** **REWARD**

```
● env.py
```

```
class InsertionEnv:
    def reset(self):
        # TODO: auto task reset
        pick_and_place(obj, target)
        go_home()
        ...

    def get_reward(self, obs, act):
        # TODO: scalar reward
        mask = sam3(obs['left'])
        pos = boundlsdf(obs, mask)
        ...

    def get_observation(self):
        ...

    def step(self, act):
        ...
```

BUILD

RUN

Human User
task objective & feedback

Coding Agent * K'

Tool APIs

PERCEPTION PLANNING CONTROL



POLICY IMPROVEMENT

01 🔍 Literature review

PLD RL-Token

CaP-X

02 💡 Propose algorithm variant

+ Heuristics

+ Off2On RL

+ Code-as-policy + BC

03 <> Optimize Infra

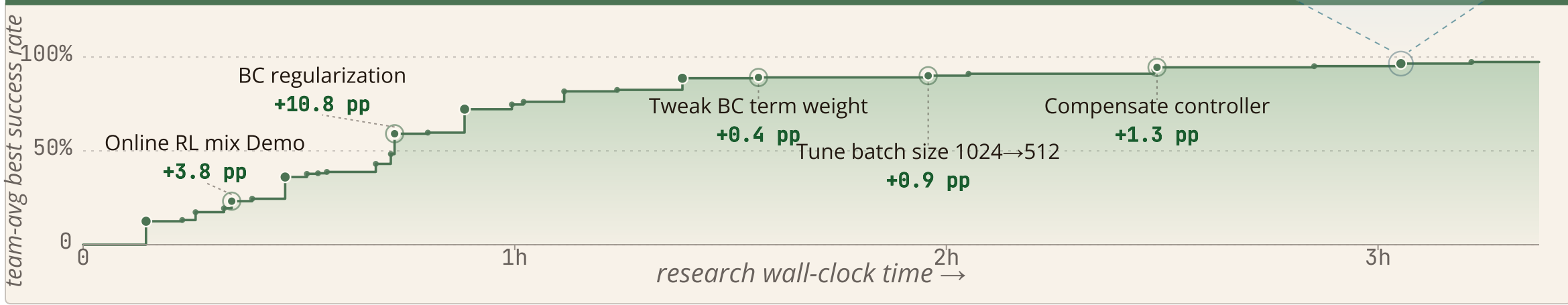
Data Sampler Param Sweep

04 📊 Summarize experiment result

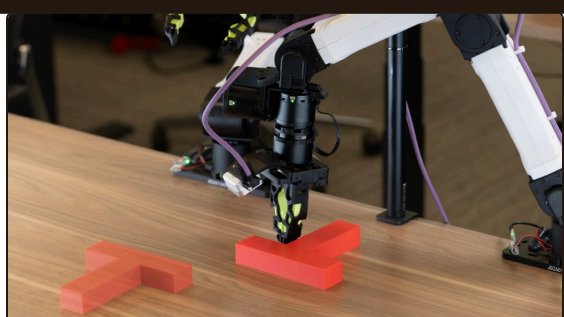
rollouts

logs

HILLCLIMB TIMELINE



REAL-WORLD TASKS



GPU insertion

Pin insertion

Push-T

Zip tie cutting