HumanoidGen: Data Generation for Bimanual Dexterous Manipulation via LLM Reasoning

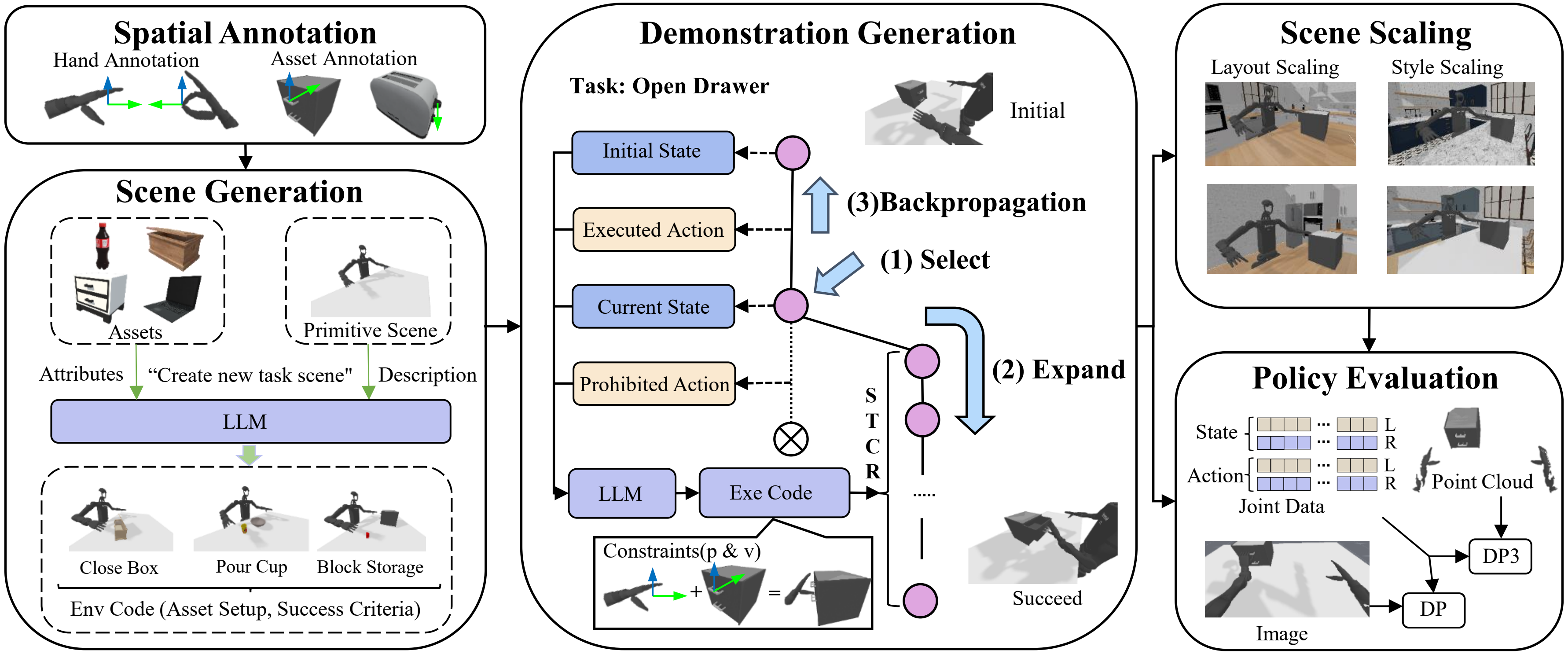
Zhi Jing[1,2] , Siyuan Yang[3,2] , Jicong Ao[2] , Ting Xiao[4] ,
Yu-Gang Jiang[1] , Chenjia Bai[†2]

[1]Fudan University[*], [2] Institute of Artificial Intelligence (TeleAI), China Telecom[*],
[3]University of Science and Technology of China,
[4]East China University of Science and Technology

[*]Equally leading organizations [†] Corresponding Author

Content

- **Motivation**
- **Method**
- **Experiments**
- **Conclusion**

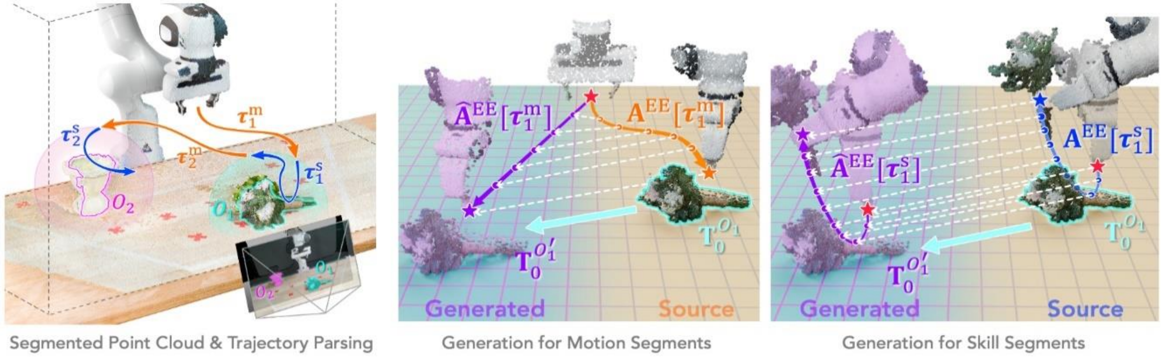


Motivation

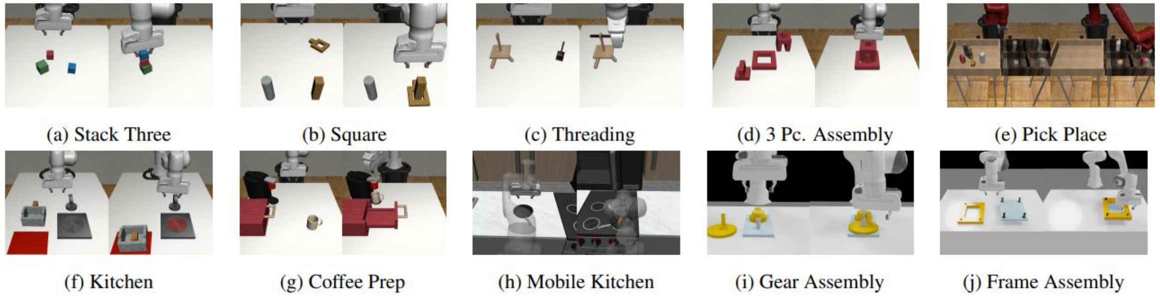
Manipulation Data Generation

Guided by Expert Demonstrations

DemoGen (RSS 2025)

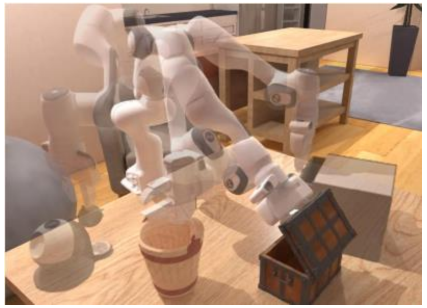


MimicGen (CoRL 2023)

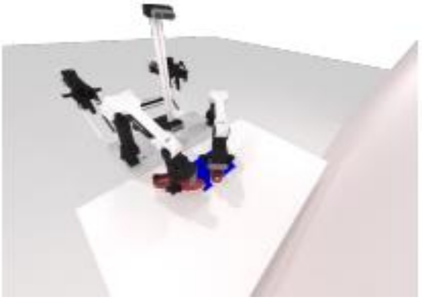


Without Expert Demonstrations

Motion Planner Based



GenSim2 (CoRL 2024)



RoboTwin (CVPR 2025)

Reinforcement Learning Based

RoboGen (ICML 2024)

Motivation

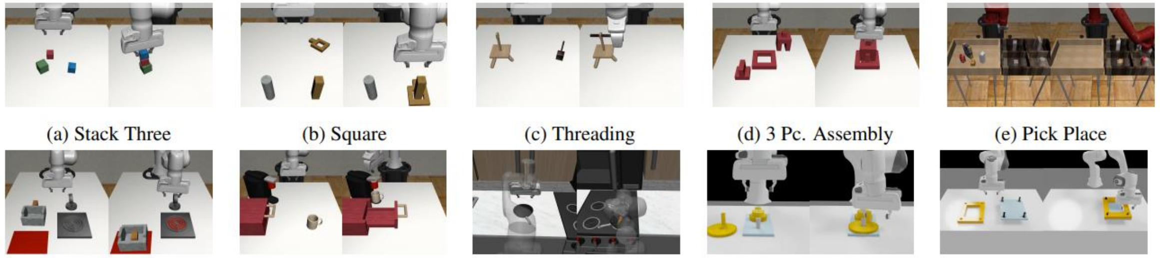
Manipulation Data Generation

Guided by Expert Demonstrations

DemoGen (RSS 2025)

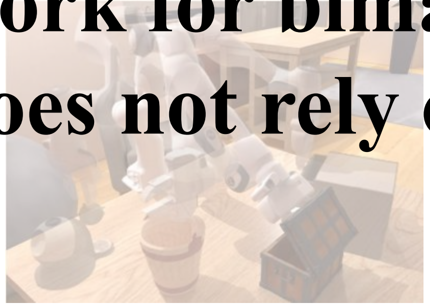
Is there a task generation framework for bimanual dexterous long-horizon manipulation that does not rely on expert demonstrations?

MimicGen (CoRL 2023)



Without Expert Demonstrations

Motion Planner Based



GenSim2 (CoRL 2024)

RoboTwin (CVPR 2025)

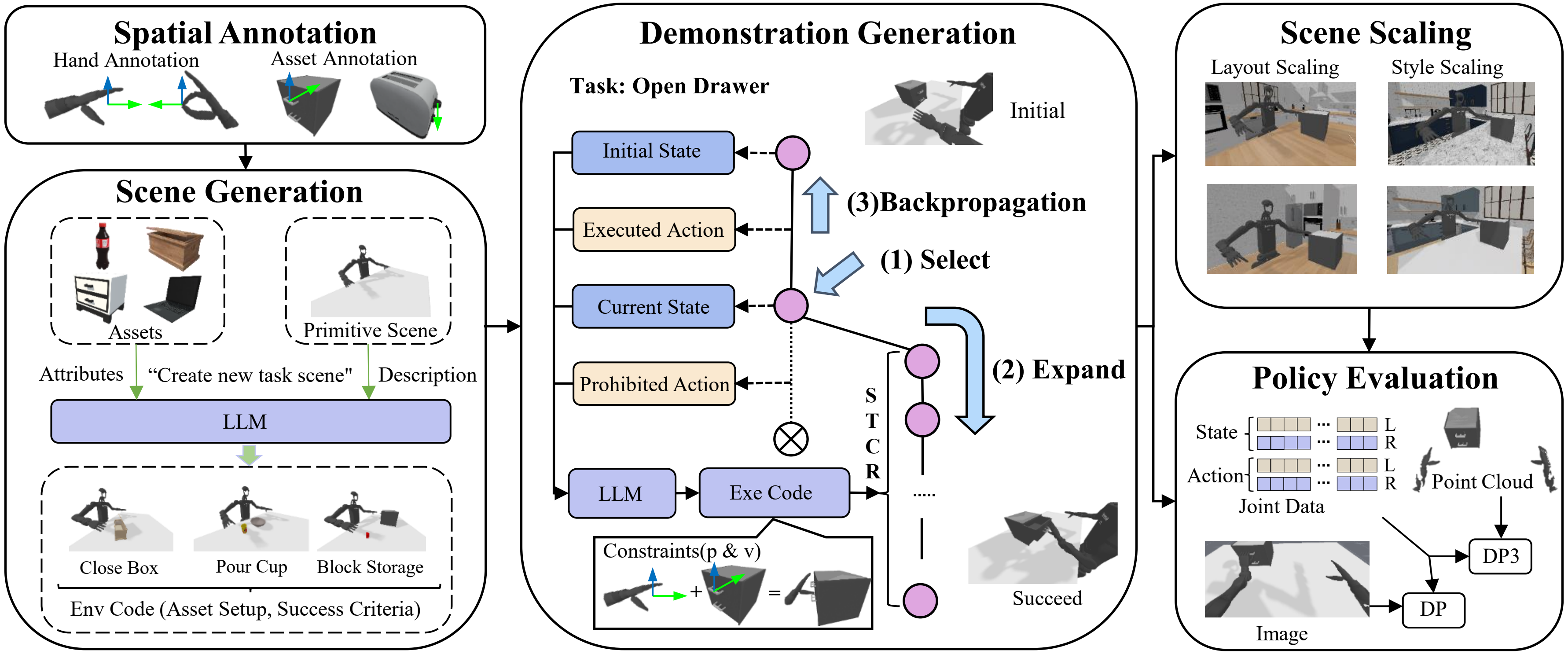
Reinforcement Learning Based

RoboGen (ICML 2024)

NEURAL INFORMATION PROCESSING SYSTEMS

Content

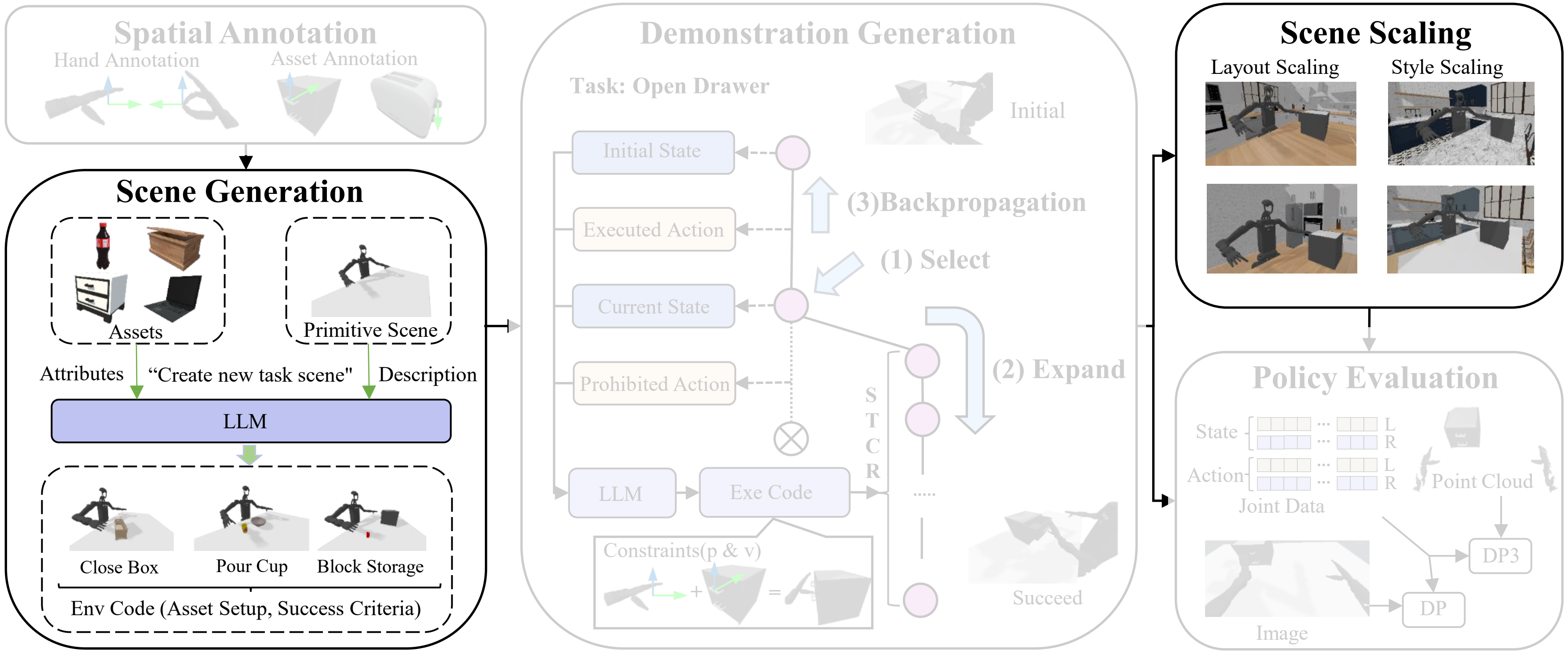
- Motivation
- **Method**
- Experiments
- Conclusion



Method

Scene Setting

Room-level Scene Scaling



Tabletop-level Scene Generation

Method

Scene Setting (Table-level)

Assets

Assets Attributes

Type Name (Box …)

Id Range (Box: 0-10)

Rigid or Articulated

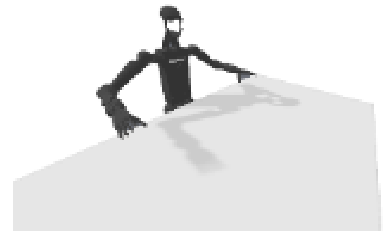
Bounding Box

Assets Description

Task Instruction:

Close the box using the right hand. (Clear)

Design a new task. (Unclear)



Scene Configuration

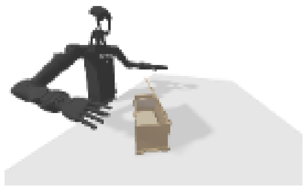
Scene Description

Robot Initial Pose

Joint Initial State

Coordinate System

Reasoning LLM



Scene Setting Code

1. Which assets are used?

_add_object(“**Box**”,**0**)

2. State definition?

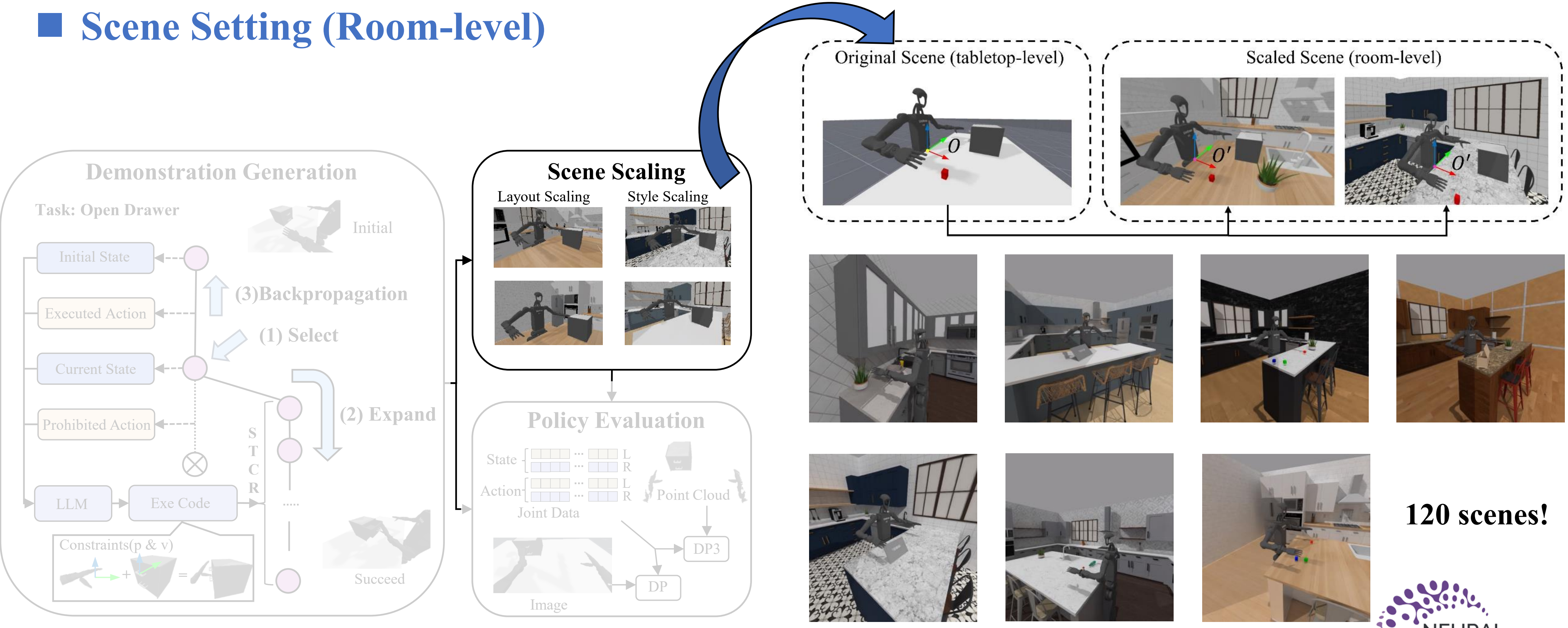
_set_object_state(**“Box”, pose=…,openness=…**)

3. Success criteria?

box.openness<close_limit (using the state obtained from the simulation environment)

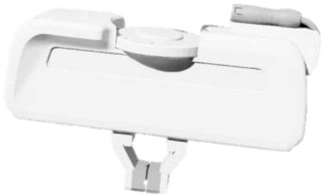
Method

Scene Setting (Room-level)



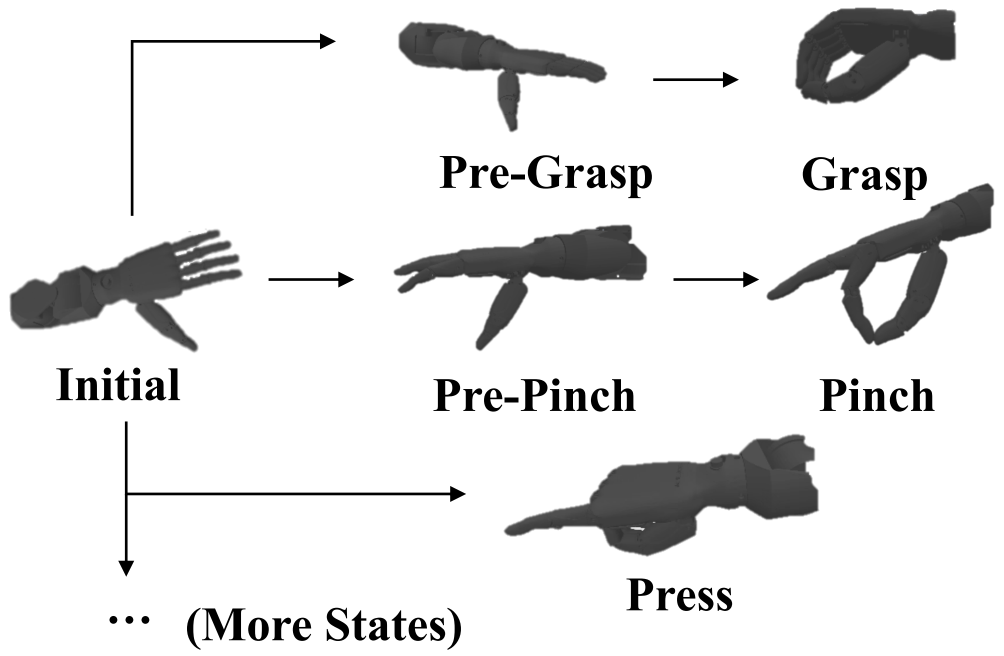
Method

Demonstration Generation



Open or Close

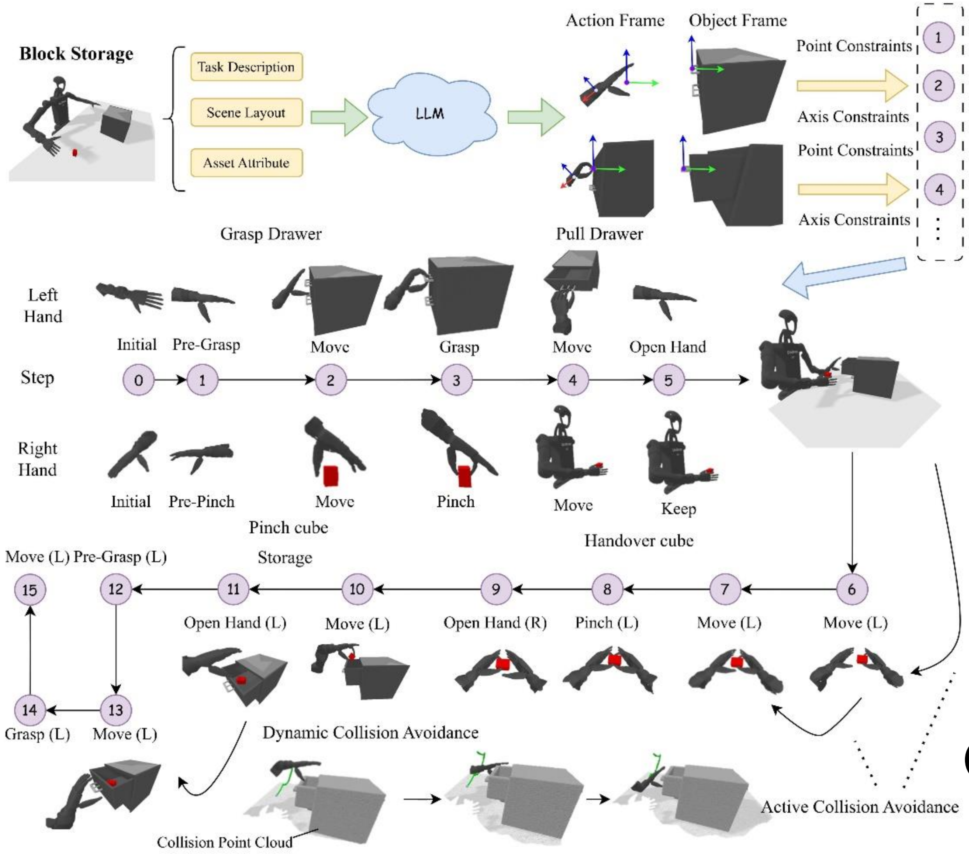
Task
- Arm Movements
- +
- Hand Operations

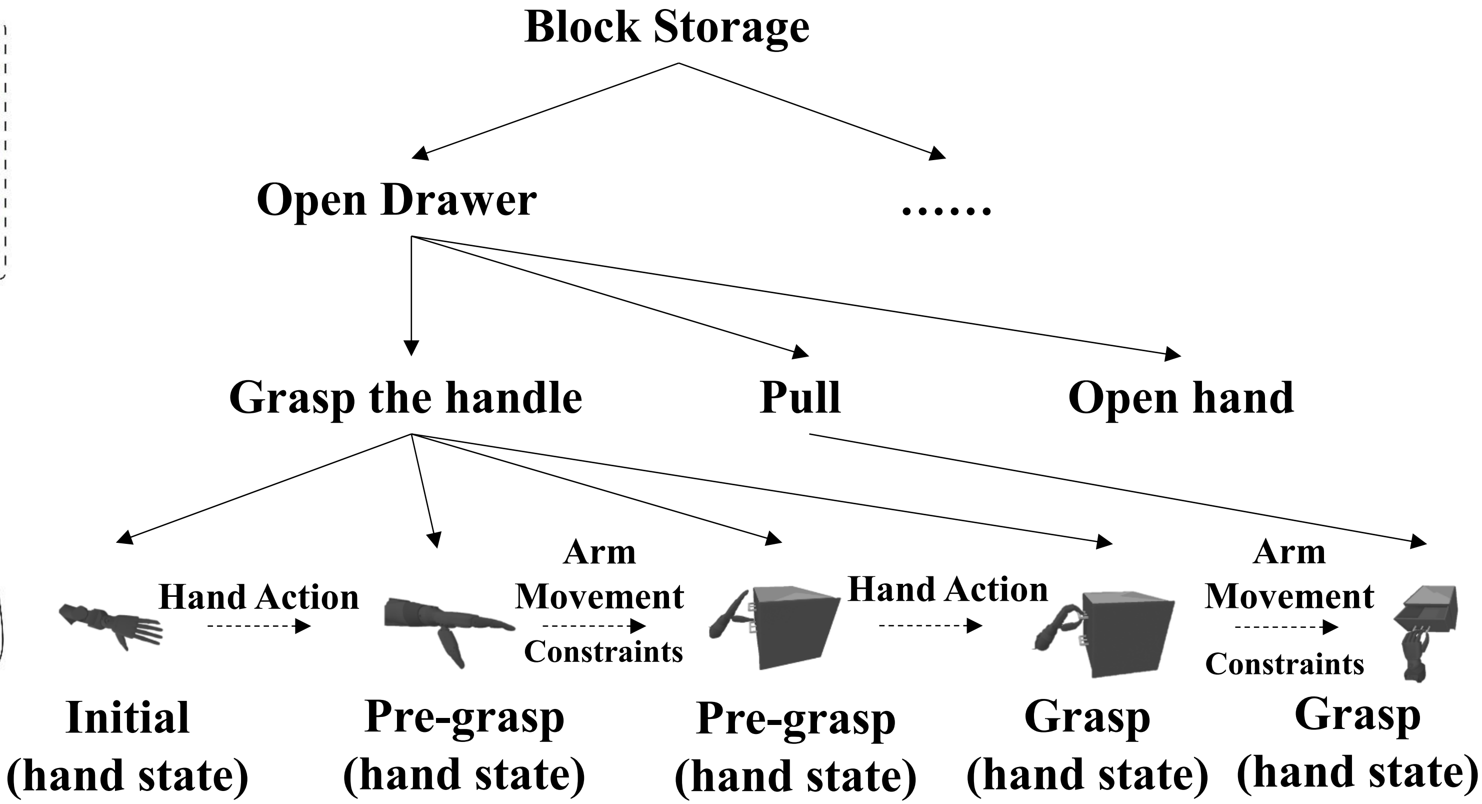


NEURAL INFORMATION PROCESSING SYSTEMS

Method

Demonstration Generation





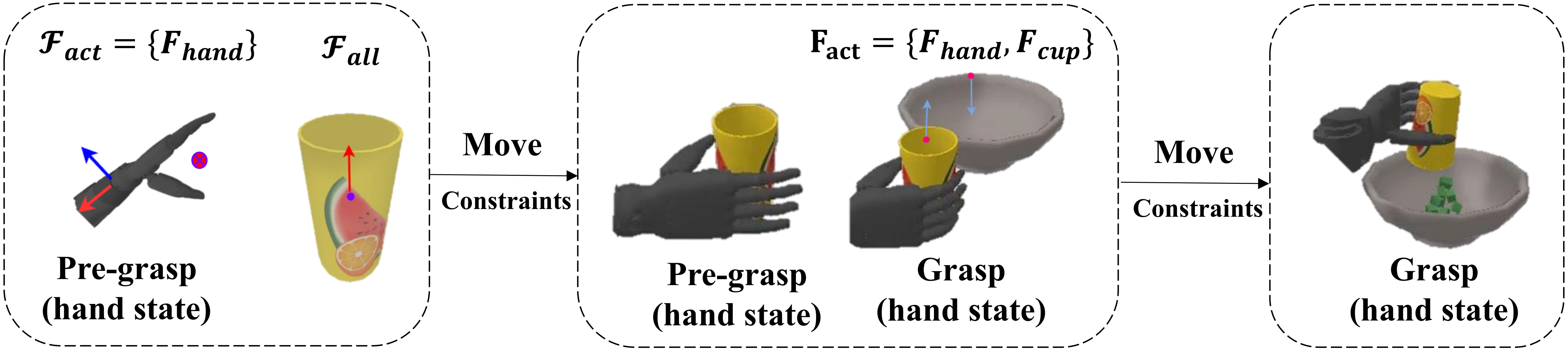
Less is more — fine-grained instructions can be composed into diverse manipulation tasks!

NEURAL INFORMATION PROCESSING SYSTEMS

Method

■ Demonstration Generation

How can we define constraints to achieve motion planning with different objectives?



Constraints

Goal Constraints:

$$\arg\min_{\theta_T} \sum_{c_i \in C^{\mathrm{goal}}} w_i c_i(u_{\mathrm{act}}, u_{\mathrm{all}}) + w_{\mathrm{reg}} \|\theta_T - \theta_{\mathrm{nominal}}\|,$$

$$\text{s.t.} \begin{cases} \mathbf{e}_T = f_{\mathrm{FK}}(\theta_T) & \text{(Kinematic constraints)} \\ c(u_{act}, u_{all}) \in [c_{\mathrm{lower}}, c_{\mathrm{upper}}], \forall c \in C^{\mathrm{goal}} & \text{(Relational action constraints)} \\ \theta_T \in \Theta_{\mathrm{collision\_free}} & \text{(Collision avoidance constraints)} \end{cases},$$

Path Constraints:

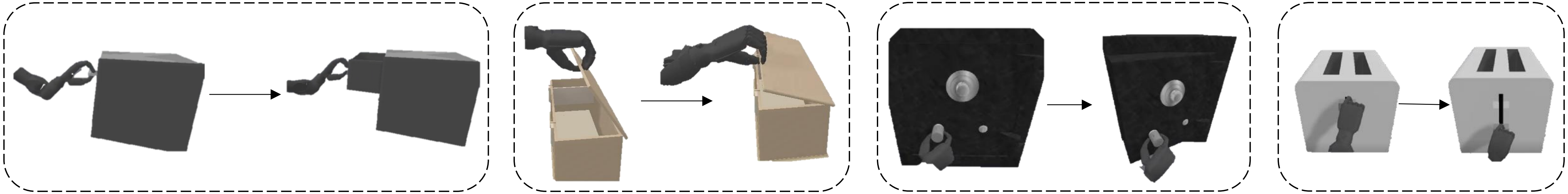
$$\arg\min_{\theta_{t\in[1,T-1]}} \mathrm{Cost}(\theta_t), \quad \text{s.t.} \begin{cases} \mathbf{e}_t = f_{\mathrm{FK}}(\theta_t), \forall t \in [1, T-1] & \text{(Kinematic constraint)} \\ c(u_{act}, u_{all}) \in [c_{\mathrm{lower}}, c_{\mathrm{upper}}], \forall c \in C^{\mathrm{path}} & \text{(Relational action constraints)}, \\ \theta_t \in \Theta_{\mathrm{collision\_free}}, \forall t \in [1, T-1] & \text{(Collision avoidance)} \end{cases}$$

NEURAL INFORMATION PROCESSING SYSTEMS

Method

■ Demonstration Generation

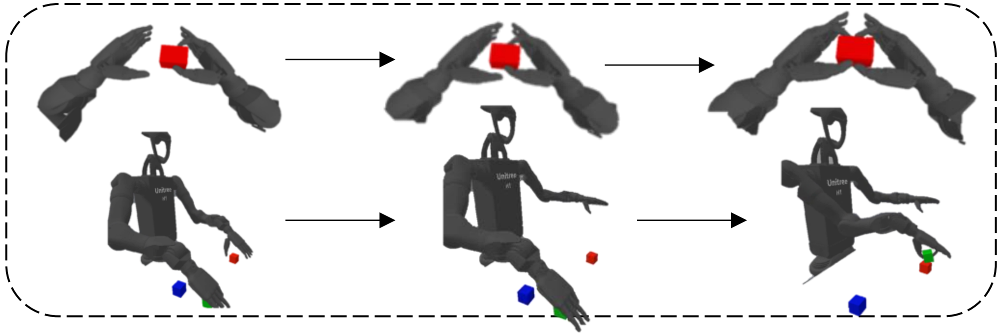
Applicable to articulated objects? Yes!



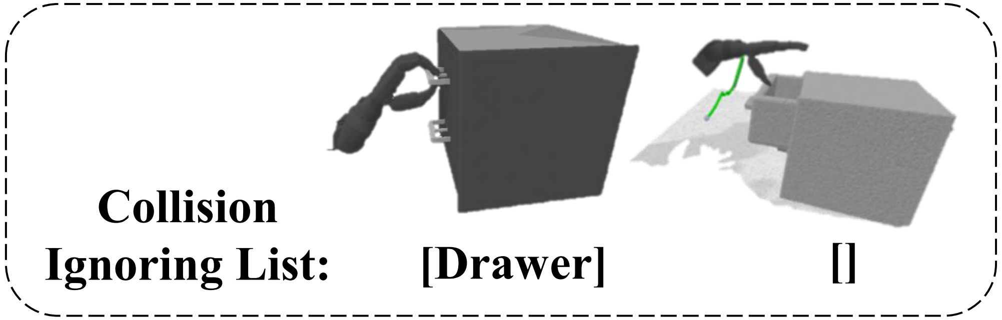
Openness: represent the degree of joint.
The extraction of key points or axes corresponding to a specific openness value is automatically implemented.

Is there collision avoidance? Yes!

Active Collision Avoidance



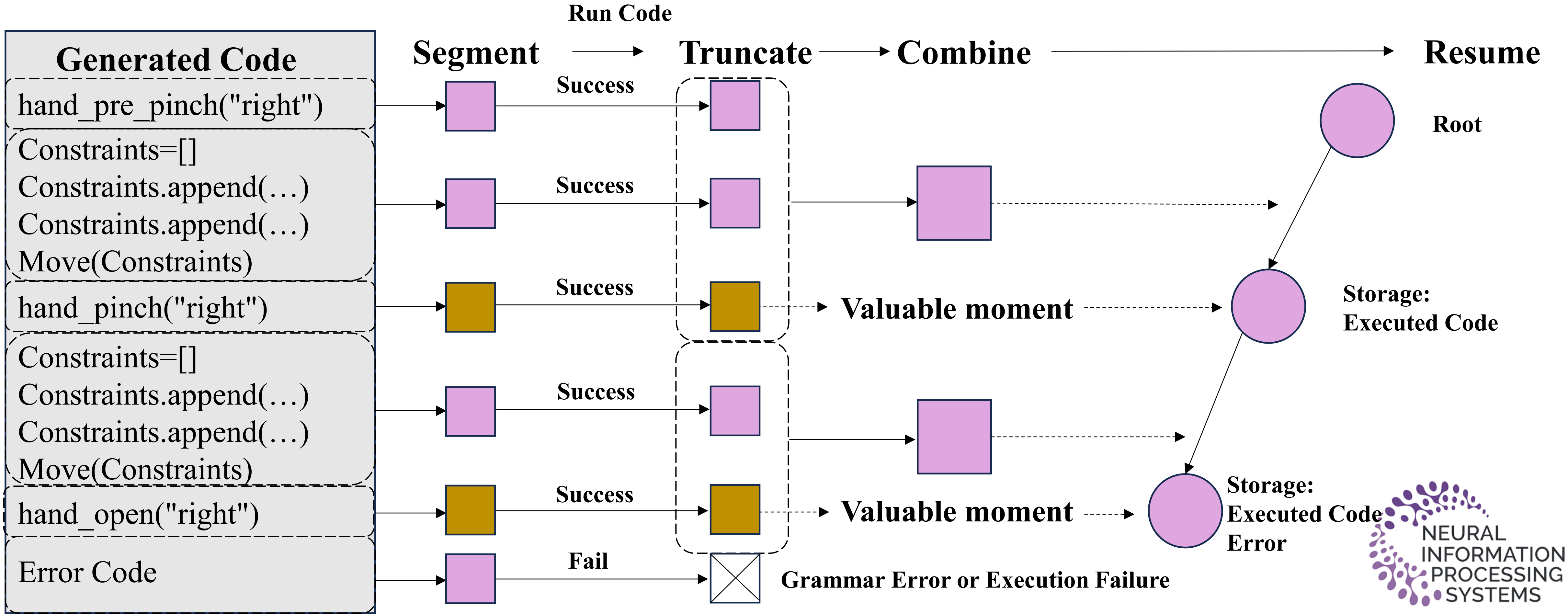
Dynamic Collision Avoidance



Method

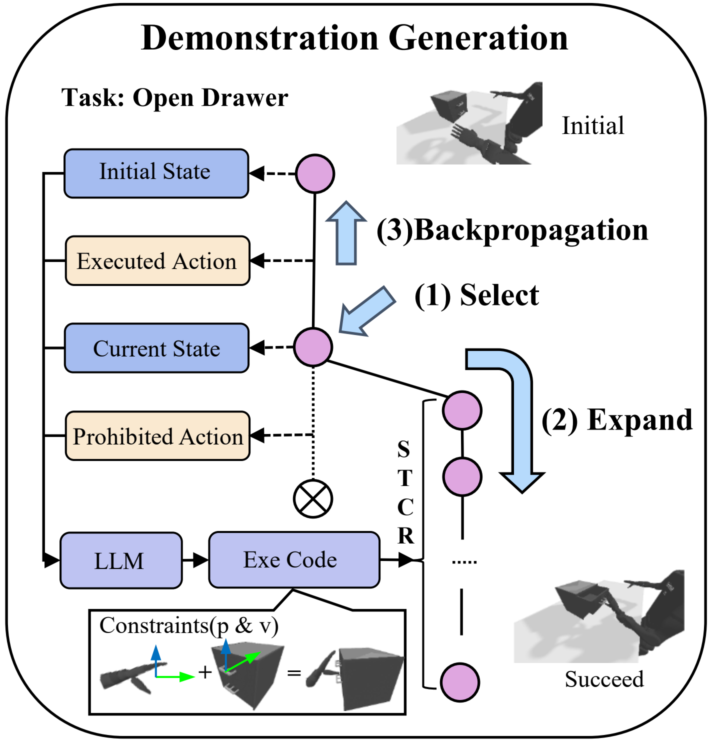
■ Demonstration Generation

STCR mechanism



Method

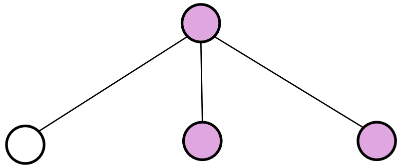
■ Demonstration Generation



(1) Selection

Starts from the root node and explores downwards to find the node to expand.

$$\mathrm{SelectPolicy}(n) = \mathop{\arg\max}_{S' \in \mathrm{Children}(n) \cup \{\varnothing\}} Q_{\mathrm{DUCB}}(n, S'),$$



(2) Expansion

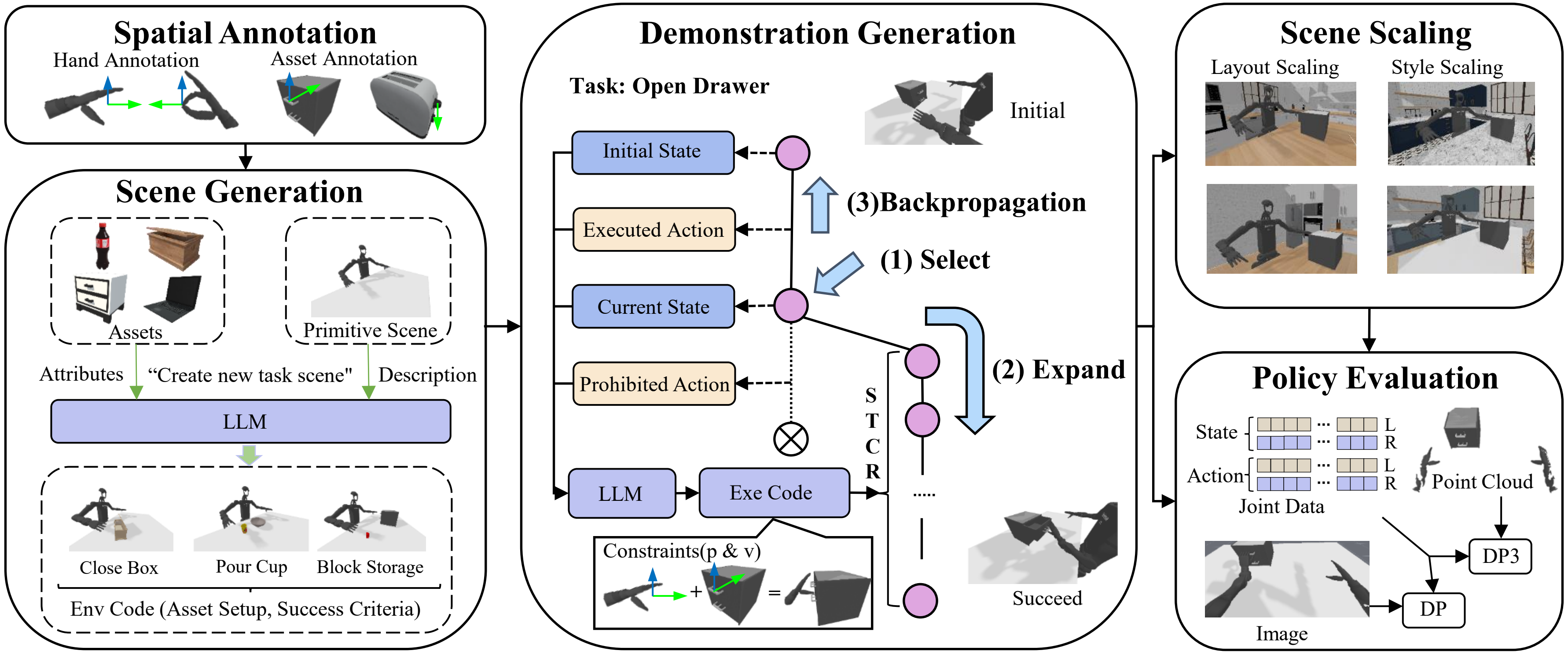
Using the information stored in the exploration node, a prompt is constructed and fed into the reasoning LLM to infer new executable code. The newly generated code will also undergo STCR processing to become a subtree.

(3) Backpropagation

Updates the visited number and reward based on intrinsic exploration values, incentivizing valuable moments such as successful grasps or stable object handling.

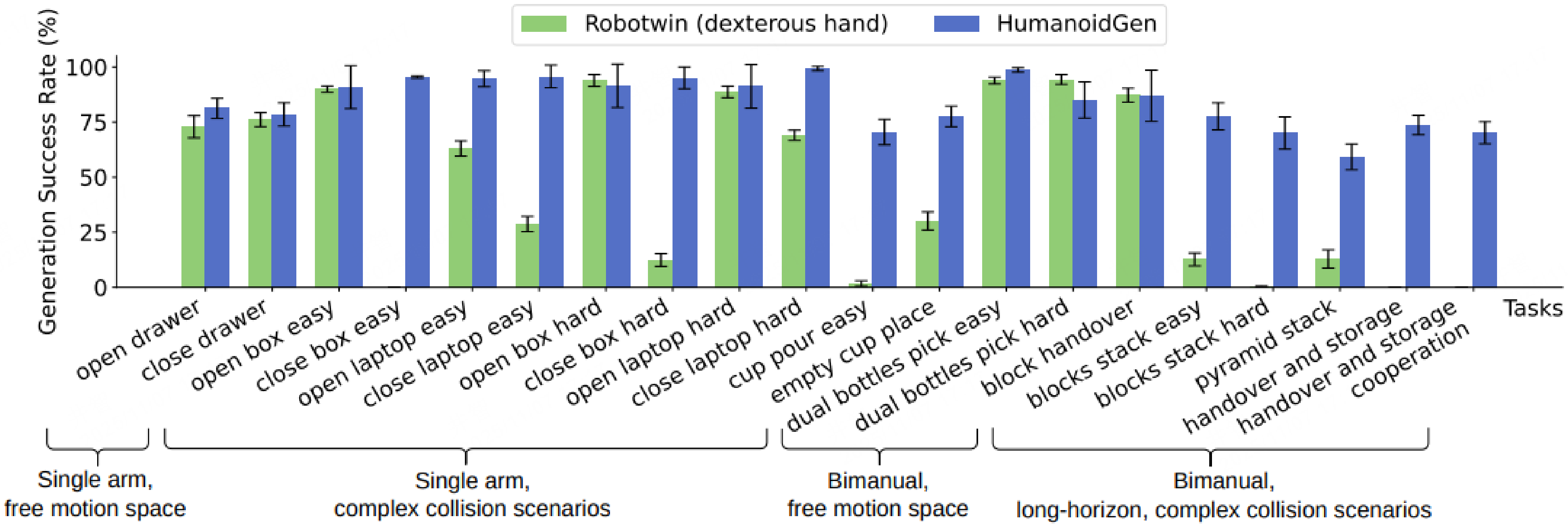
Content

- Motivation
- Method
- Experiments
- Conclusion



Experiments

Evaluation of Data Generation and Execution



Our framework outperforms RoboTwin in long-horizon and complex collision tasks, demonstrating superior dexterous manipulation performance through dynamic collision management. On average, it achieves an 81.37% higher success rate, while Robotwin fails in most long-horizon, highly collisional tasks.

Experiments

■ Effectiveness Evaluation of MCTS

Method	Success rate (%)	Token consumption (K)
Block Stack Single		
Non-MCTS	63.3 ± 6.24	15.3 ± 1.90
$\text{MCTS}_{N=2}$	98.3 ± 2.36	19.3 ± 7.04
Blocks Stack Easy		
Non-MCTS	46.7 ± 2.36	14.8 ± 1.64
$\text{MCTS}_{N=2}$	83.3 ± 5.56	21.6 ± 7.78
$\text{MCTS}_{N=3}$	95.0 ± 4.08	22.8 ± 9.13
Blocks Stack Hard		
Non-MCTS	18.3 ± 6.24	16.0 ± 0.92
$\text{MCTS}_{N=8}$	78.3 ± 2.36	69.9 ± 39.24
$\text{MCTS}_{N=12}$	98.3 ± 2.36	78.3 ± 51.75
Pyramid Stack		
Non-MCTS	13.3 ± 6.24	16.2 ± 1.29
$\text{MCTS}_{N=8}$	76.7 ± 4.71	80.0 ± 31.72
$\text{MCTS}_{N=12}$	90.0 ± 4.08	89.6 ± 44.61

Table 1: The evaluation results of applying different numbers of max MCTS exploration steps N and non-MCTS in four tasks.

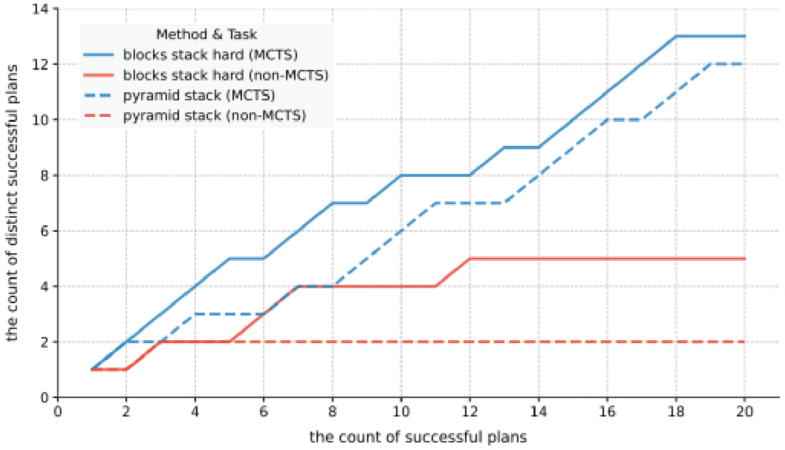
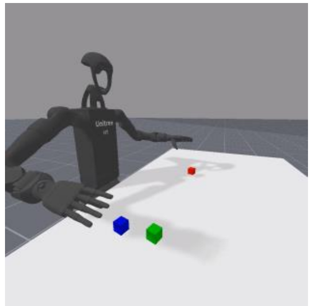
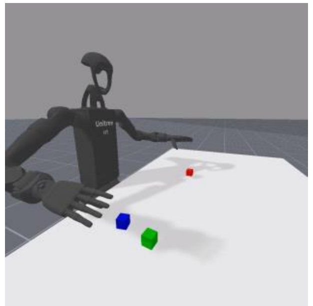
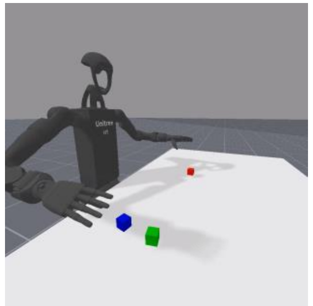
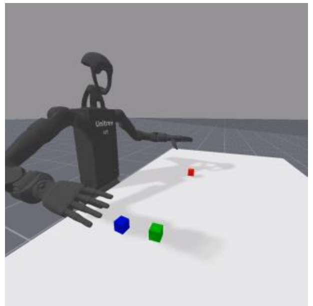


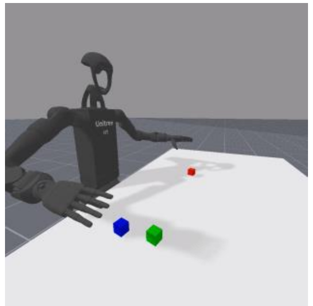
Figure 6: The variation of the count of distinct successful plans for MCTS and non-MCTS with the count of successful plans.











MCTS improves reasoning success and plan diversity with minimal extra tokens, enabling error correction and diverse strategy exploration in complex stacking tasks.

Experiments

HGen-Bench Evaluation

Table 3: We present the DP and DP3 results for all 20 tasks using 100, 50, and 20 trajectories generated by our method, and evaluate the success rates across 14 tasks using 3 random seeds.

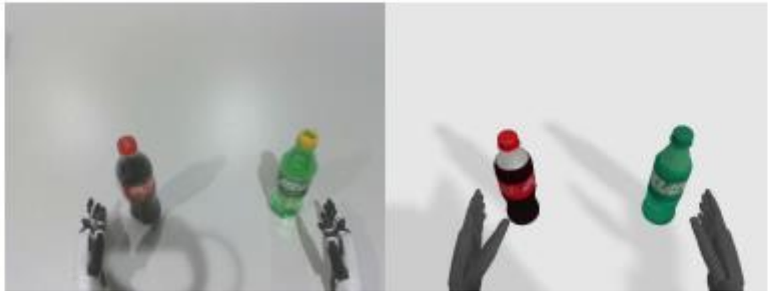
Num of Demonstrations	20	50	100		20	50	100
Blocks Stack Easy				**Close Drawer**			
DP3	0.0±0.0	0.0±0.0	22.8±16.5	DP3	83.3±17.6	94.4±7.9	92.6±8.3
DP	0.0±0.0	0.0±0.0	0.0±0.0	DP	95.6±3.7	100.0±0.0	100.0±0.0
Cup Pour Easy				**Dual Bottles Pick Easy**			
DP3	67.8±10.8	75.6±9.6	72.2±7.9	DP3	75.9±17.8	96.3±6.9	93.9±7.6
DP	0.0±0.0	2.2±6.3	0.0±0.0	DP	0.0±0.0	0.0±0.0	0.0±0.0
Dual Bottles Pick Hard				**Empty Cup Place**			
DP3	88.9±13.6	90.7±11.4	94.4±7.9	DP3	25.0±8.2	18.3±4.7	33.3±7.1
DP	0.0±0.0	0.0±0.0	0.0±0.0	DP	0.0±0.0	0.0±0.0	6.7±13.3
Open Box Easy				**Open Box Hard**			
DP3	85.6±8.0	95.6±4.4	95.0±4.1	DP3	95.6±5.5	96.1±4.6	98.3±3.3
DP	93.3±13.3	100.0±0.0	100.0±0.0	DP	11.1±19.1	93.3±9.4	100.0±0.0
Open Drawer				**Open Laptop Hard**			
DP3	58.3±8.3	76.0±13.1	84.4±11.3	DP3	100.0±0.0	100.0±0.0	100.0±0.0
DP	17.8±22.0	13.3±18.9	48.9±31.4	DP	15.6±22.7	11.1±9.9	35.6±32.4
Close Box Hard				**Close Laptop Easy**			
DP3	88.9±17.6	96.3±6.9	96.3±6.9	DP3	100.0±0.0	100.0±0.0	100.0±0.0
DP	*82.2±22.0	*51.1±19.1	31.1±28.5	DP	37.8±23.9	40.0±23.1	48.9±25.1
Handover and Storage				**Blocks Stack Hard**			
DP3	0.0±0.0	0.0±0.0	0.0±0.0	DP3	0.0±0.0	0.0±0.0	0.0±0.0
DP	0.0±0.0	0.0±0.0	0.0±0.0	DP	0.0±0.0	0.0±0.0	0.0±0.0
Block Handover				**Close Box Easy**			
DP3	0.0±0.0	0.0±0.0	0.0±0.0	DP3	100.0±0.0	98.3±3.3	99.4±1.6
DP	0.0±0.0	0.0±0.0	0.0±0.0	DP	97.8±6.3	100.0±0.0	91.1±13.7
Close Laptop Hard				**Handover and Storage Cooperation**			
DP3	92.6±8.3	94.4±7.9	96.3±6.9	DP3	0.0±0.0	0.0±0.0	0.0±0.0
DP	*46.7±13.3	*42.2±34.6	33.3±26.7	DP	0.0±0.0	0.0±0.0	0.0±0.0
Open Laptop Easy				**Pyramid Stack**			
DP3	71.1±5.7	77.2±7.5	81.1±9.4	DP3	0.0±0.0	0.0±0.0	0.0±0.0
DP	*75.6±18.3	*71.1±16.6	60.0±16.3	DP	0.0±0.0	0.0±0.0	0.0±0.0

We provide evaluations of DP and DP3 on our collected dataset.

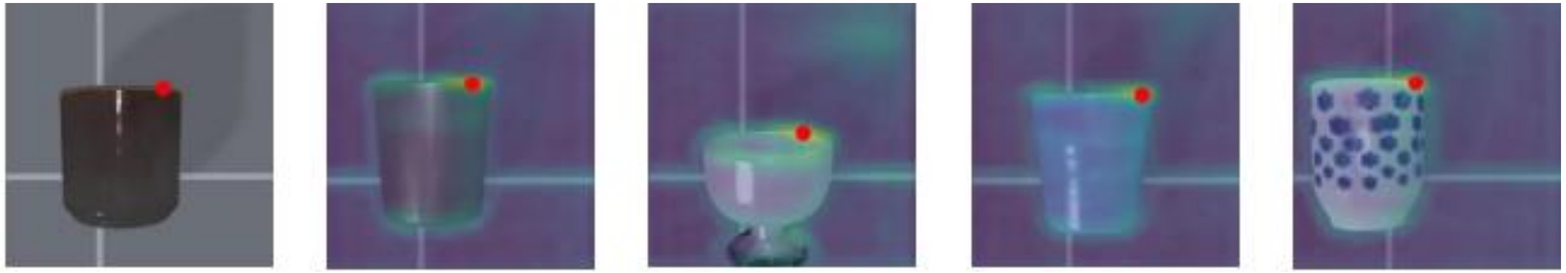
Experiments

Other experiments

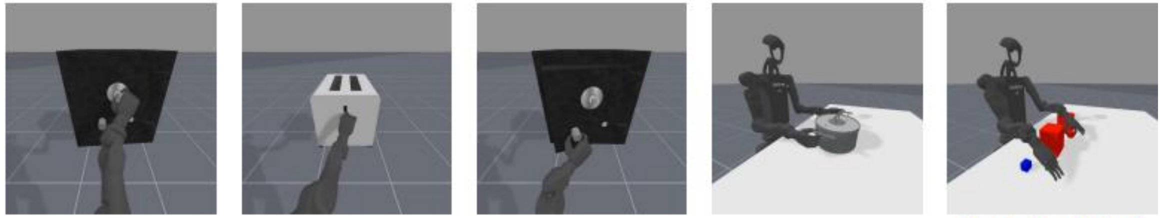
Real-World Experiments



Automatic Asset Annotation



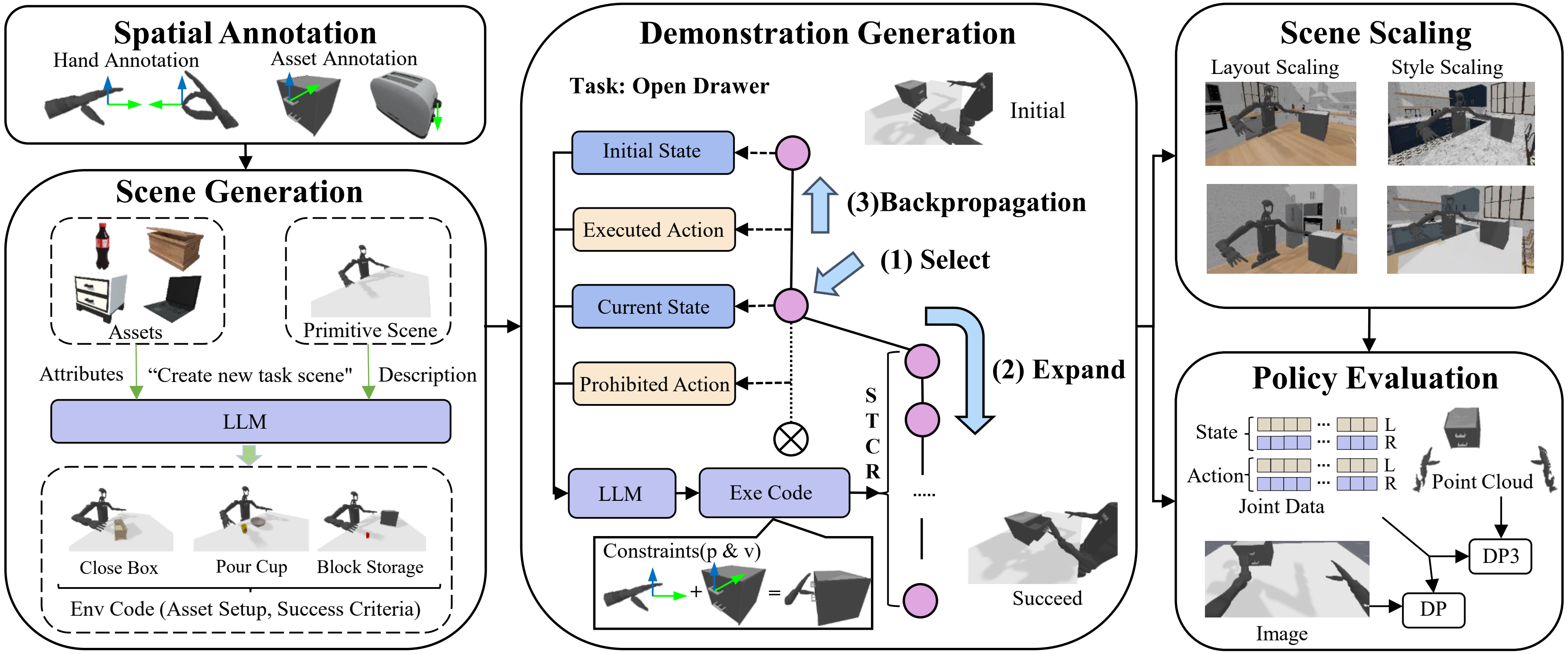
Additional Challenging Manipulation Tasks



Rotate Safe Knob | Press Toaster | Open Safe Door | Dual Lift Pot | Blocks Stack Hard With Barrier

Content

- Motivation
- Method
- Experiments
- Conclusion



Conclusion

We propose a complete data generation framework for bimanual dexterous manipulation.

- Home Page: https://openhumanoidgen.github.io/
- Code Project: https://github.com/TeleHuman/HumanoidGen
- Connect Me: jingzhi2021@qq.com