Towards Reliable Code-as-Policies: A Neuro-Symbolic Framework for Embodied Task Planning

Sanghyun Ahn, Wonje Choi, Junyong Lee, Jinwoo Park, Honguk Woo

Department of Computer Science and Engineering, Sungkyunkwan University

{shyuni5, wjchoi1995, ljy7488, pjw971022, hwoo}@skku.edu

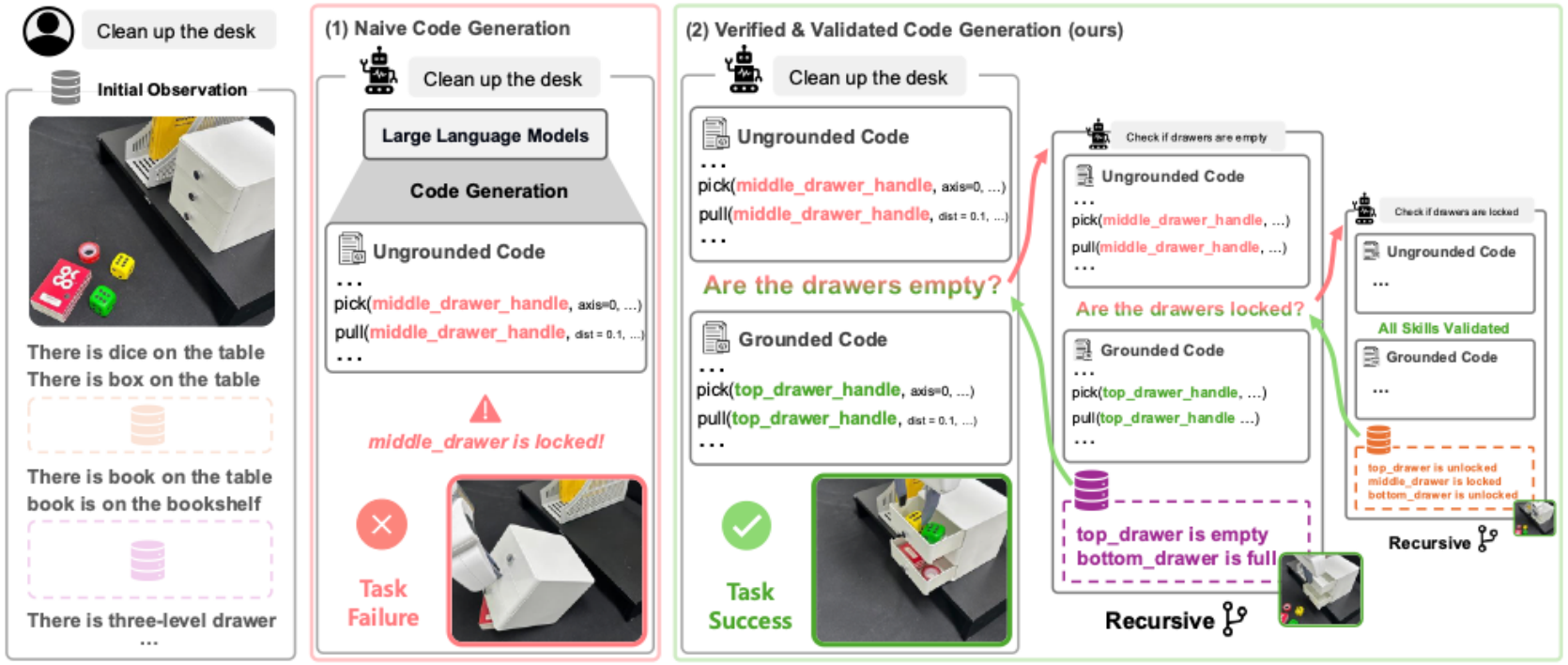
NEURAL INFORMATION PROCESSING SYSTEMS

SUNG KYUN KWAN UNIVERSITY(SKKU)

NEURAL INFORMATION PROCESSING SYSTEMS

SUNG KYUN KWAN UNIVERSITY(SKKU)

Introduction

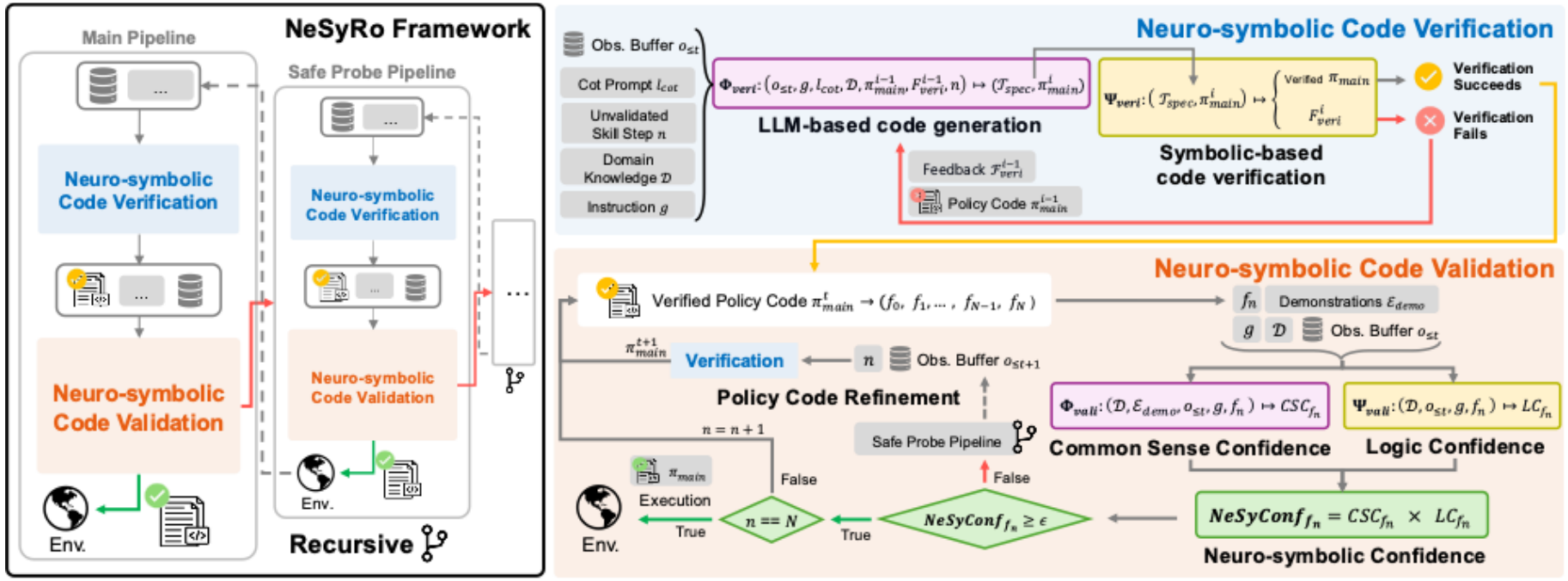


- Code-as-policies for embodied control
  - LLMs generate executable Python code for robot task planning
  - Suffer from limited environmental grounding in dynamic, partially observable environments
  - **Example failure**: Attempting to open locked drawer without checking → irreversible action

NEURAL INFORMATION PROCESSING SYSTEMS

SUNG KYUN KWAN UNIVERSITY(SKKU)

NeSyRo(Neuro-Symbolic Robot task planning) Framework



- Recursive Neuro-symbolic code Verification & Validation

- **Neuro-symbolic code verification**: Ensures generated code satisfies task specification
- **Neuro-symbolic code validation**: Assesses skill executability and triggers safe probes when confidence is low
- Recursive composition of verification and validation until all skills are grounded

Evaluation: Setting

- Environment

- For simulation: RLBench with 7-DoF Franka Emika robot
- For real-world: Franka Emika Research 3
- Four observability levels: High/Low/Stochastic/Complete incompleteness

- Task Types

Task Type	Probe Type
Object Relocation (e.g., moving tomatoes on a plate)	Robot Pose Adjust (e.g., verifying which item is a tomato)
Object Interaction (e.g., opening a drawer)	Object State Check (e.g., checking whether a drawer is locked)
Auxiliary Manipulation (e.g., opening a drawer in a dark room)	Object State Change (e.g., turning on the light to locate the drawer)
Long-Horizon (e.g., placing a tomato inside a drawer)	Uses two or more of the above probe types depending on task structure and uncertainty

Evaluation: Setting

- Baselines
  - Code-as-policies methods: CaP, CaP w/ Lemur, CaP w/ CodeSift
  - LLM-based replanning: LLM-Planner, AutoGen

- Metrics
  - Success Rate (SR) (%): percentage of tasks fully completed
  - Goal Condition (GC) (%): percentage of sub-goals achieved
  - Irreversible Actions (IA): count of unsafe actions in real-world

NEURAL INFORMATION PROCESSING SYSTEMS

SUNG KYUN KWAN UNIVERSITY(SKKU)

Evaluation: Performance

Methods	*High*		*Low*		*Stochastic*		*Complete*	
	SR	GC	SR	GC	SR	GC	SR	GC
Task Type: Object Relocation								
CaP	25.0±7.1	41.5±8.8	30.0±0.0	43.8±1.8	10.0±0.0	36.3±1.8	90.0±0.0	92.5±3.5
CaP w/ Lemur	25.0±7.1	43.8±5.3	30.0±0.0	43.8±1.8	10.0±0.0	36.3±1.8	90.0±0.0	96.3±1.8
CaP w/ CodeSift	55.0±7.1	72.5±3.5	50.0±0.0	57.5±3.5	40.0±0.0	52.5±3.5	95.0±7.1	95.0±7.1
LLM-Planner	30.0±0.0	35.0±7.1	50.0±0.0	58.8±5.3	30.0±0.0	43.8±5.3	80.0±0.0	88.8±5.3
AutoGen	30.0±0.0	35.0±7.1	55.0±7.1	60.0±7.1	40.0±14.1	47.5±10.6	85.0±7.1	87.5±10.6
NeSyRo	70.0±14.1	72.5±10.6	75.0±7.1	87.5±3.5	65.0±7.1	75.0±0.0	95.0±7.1	97.5±3.5
Task Type: Object Interaction								
CaP	20.0±14.1	35.0±7.1	25.0±7.1	40.0±3.5	35.0±7.1	51.3±5.3	75.0±7.1	77.5±7.1
CaP w/ Lemur	35.0±7.1	47.5±7.1	35.0±7.1	47.5±3.5	30.0±14.1	46.3±12.1	85.0±7.1	86.3±8.8
CaP w/ CodeSift	40.0±0.0	65.0±7.1	50.0±14.1	55.0±7.1	40.0±0.0	60.0±14.1	90.0±14.1	90.0±14.1
LLM-Planner	5.0±7.1	15.0±0.0	40.0±14.1	53.8±5.3	35.0±7.1	42.5±14.1	55.0±7.1	63.8±8.8
AutoGen	40.0±0.0	48.8±1.8	50.0±0.0	58.8±5.3	50.0±0.0	57.5±0.0	75.0±7.1	76.3±8.8
NeSyRo	70.0±0.0	76.3±1.8	80.0±0.0	83.8±1.8	70.0±14.1	73.8±8.8	90.0±0.0	92.5±0.0
Task Type: Auxiliary Manipulation								
CaP	25.0±7.1	25.0±7.1	50.0±0.0	51.3±1.8	40.0±0.0	45.8±8.3	85.0±7.1	90.0±4.7
CaP w/ Lemur	30.0±14.1	30.0±14.1	50.0±14.1	58.3±7.1	30.0±14.1	34.2±15.3	85.0±7.1	90.8±3.5
CaP w/ CodeSift	5.0±7.1	5.0±7.1	55.0±7.1	57.5±3.5	35.0±7.1	35.0±7.1	90.0±0.0	93.3±0.0
LLM-Planner	15.0±7.1	15.0±7.1	30.0±0.0	37.5±3.5	10.0±14.1	15.0±7.1	75.0±7.1	80.0±0.0
AutoGen	15.0±7.1	15.0±7.1	35.0±7.1	40.0±0.0	20.0±0.0	22.5±3.5	80.0±0.0	80.0±0.0
NeSyRo	60.0±0.0	80.8±1.2	70.0±14.1	74.2±13.0	70.0±14.1	85.8±5.9	95.0±7.1	96.7±4.7
Task Type: Long-Horizon								
CaP	0.0±0.0	0.0±0.0	20.0±0.0	40.4±6.6	0.0±0.0	0.7±1.0	40.0±14.1	53.8±3.4
CaP w/ Lemur	0.0±0.0	0.0±0.0	30.0±0.0	47.1±0.0	0.0±0.0	1.6±0.2	55.0±7.1	67.1±5.1
CaP w/ CodeSift	0.0±0.0	0.0±0.0	30.0±14.1	45.8±7.9	5.0±7.1	5.0±7.1	65.0±7.1	71.4±6.7
LLM-Planner	0.0±0.0	0.0±0.0	10.0±0.0	11.4±0.0	5.0±0.0	12.9±8.1	35.0±7.1	44.4±9.9
AutoGen	0.0±0.0	5.5±3.0	30.0±14.1	39.2±10.3	20.0±0.0	28.5±7.2	50.0±0.0	55.1±0.8
NeSyRo	45.0±7.1	65.2±6.7	45.0±7.1	58.1±6.1	35.0±7.1	41.9±8.1	65.0±7.1	73.7±8.3

Table2. Performance evaluation on RLBench simulation

- Experiment results

- NeSyRo consistently outperforms baselines across all observability levels

- Under high incompleteness, NeSyRo achieves **70%** success rate compared to **25-55%** for baselines

- For long-horizon tasks, NeSyRo reaches **45%** success rate while baselines achieve only **0-30%**

NEURAL INFORMATION PROCESSING SYSTEMS

SUNG KYUN KWAN UNIVERSITY(SKKU)

Evaluation: Performance

Real-World	CaP			CaP w/ CodeSift			NeSyRo			NeSyRo-Complete		
Task Type	SR (↑)	GC (↑)	IA (↓)	SR (↑)	GC (↑)	IA (↓)	SR (↑)	GC (↑)	IA (↓)	SR (↑)	GC (↑)	IA (↓)
Object Relocation	7.5±3.5	11.3±1.8	19	12.5±3.5	19.4±4.4	4	82.5±3.5	83.8±3.5	2	85.0±7.1	90.0±3.5	2
Object Interaction	30.0±7.1	37.5±7.1	12	20.0±7.1	24.4±9.7	4	75.0±14.1	77.5±17.7	0	90.0±14.1	90.0±14.1	0
Auxiliary Manipulation	0.0±0.0	0.0±0.0	4	2.5±3.5	8.3±4.7	5	20.0±0.0	20.0±0.0	2	20.0±14.1	20.0±14.1	2
Long-Horizon	5.0±0.0	14.2±3.5	18	7.5±10.6	13.1±7.4	16	52.5±3.5	54.2±3.5	3	60.0±0.0	65.8±8.3	2
Total	10.6±0.9	15.7±0.4	53	10.6±4.4	16.3±4.4	29	57.5±3.5	58.9±4.4	7	68.8±5.3	71.5±6.5	6

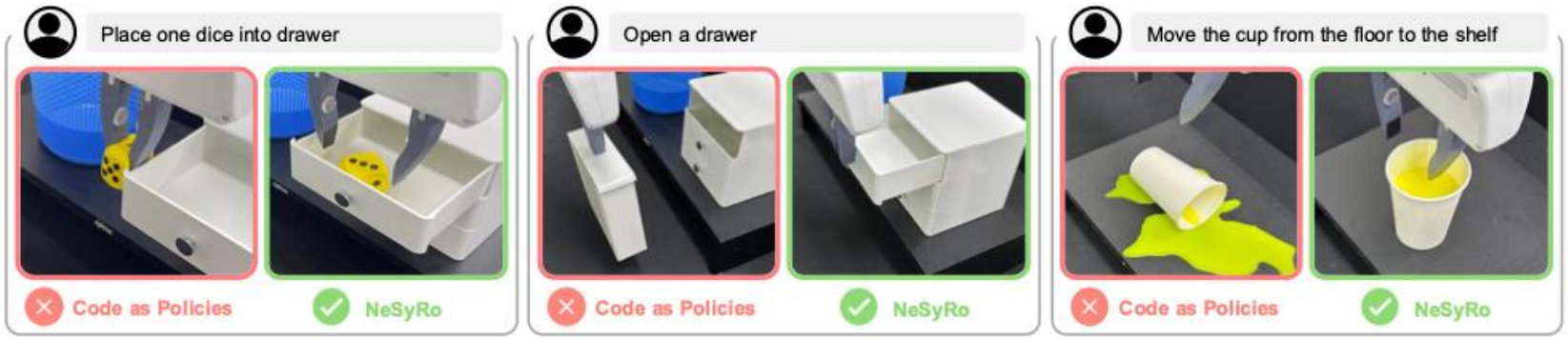


Table3. Performance evaluation on Real-World

- Experiment results

- NeSyRo achieves **57.5%** success rate compared to **10.6%** for baselines under partial observability
- NeSyRo significantly reduces irreversible actions: **7** vs **29-53** for baselines

Conclusion

- Key Contribution

- NeSyRo achieves reliable code generation through recursive neuro-symbolic verification and validation, improving success rate by **46.2%** and reducing irreversible actions by **86.8%.**

- Future Work

- Extending to probabilistic reasoning and domain knowledge-free validation