

# LogiCity: Advancing Neuro-Symbolic AI with Abstract Urban Simulation

Bowen Li<sup>1</sup> Zhaoyu Li<sup>2</sup> Qiwei Du<sup>3</sup> Jinqi Luo<sup>4</sup> Wenshan Wang<sup>1</sup> Yaqi Xie<sup>1</sup> Simon Stepputtis<sup>1</sup>  
Chen Wang<sup>3</sup> Katia Sycara<sup>1</sup> Pradeep Ravikumar<sup>1</sup> Alexander Gray<sup>5</sup> Xujie Si<sup>2</sup> Sebastian Scherer<sup>1,\*</sup>

<sup>1</sup>Carnegie Mellon University, <sup>2</sup>University of Toronto, <sup>3</sup>University at Buffalo  
<sup>4</sup>University of Pennsylvania, <sup>5</sup>Centaur AI Institute

NeurIPS 2024 D&B Track (w/ audio)

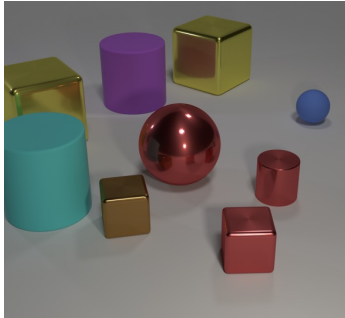


# LogiCity: Advancing Neuro-Symbolic AI with Abstract Urban Simulation

Other Neuro-Symbolic Testbeds

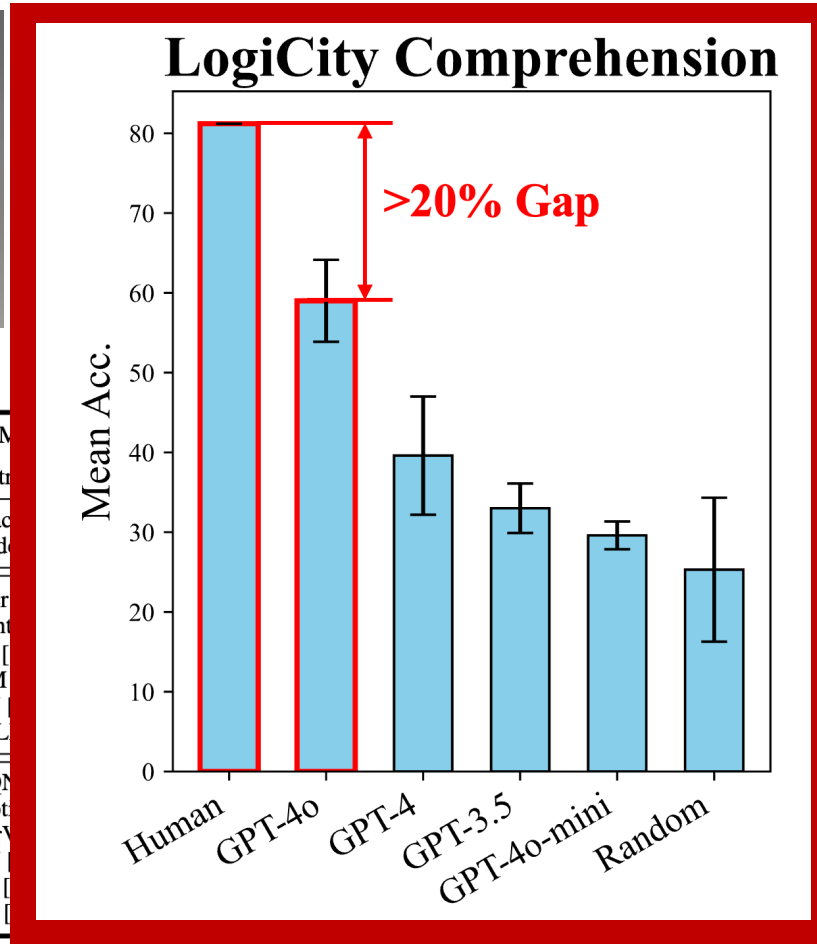
0	6	2	1	0	7	0	8	0
0	3	0	0	0	8	2	5	0
8	0	0	0	0	4	0	0	0
0	0	0	0	8	0	7	0	0
4	9	1	0	6	0	0	2	8
5	0	0	3	4	0	1	0	0
0	0	3	0	7	9	0	1	0
1	7	0	0	0	0	5	0	0
0	5	0	0	0	0	9	6	0

Visual Sudoku<sup>[1]</sup>



CLEVR<sup>[2]</sup>

Supervision	Model
	Metri
N/A	Orac
	Rand
Behavior Cloning	Popper
	MaxSynt
	HRI [
	NLM
	GNN
	ML
Reinforcement Learning	NLM-DQN
	MB-shoot
	Dreamer
	DQN
	PPO [
	A2C [



LogiCity



[1] SATNet: Bridging Deep Learning and Logical Reasoning using a Differentiable Satisfiability Solver. In. ICML 2019  
 [2] CLEVR: A Diagnostic Dataset for Compositional Language and Elementary Visual Reasoning. In. CVPR 2017  
 [3] MiniGrid & MiniWorld: Modular & Customizable Reinforcement Learning Environments for Goal-Oriented Tasks. In. NeurIPS 2023  
 [4] An Environment for Autonomous Driving Decision-Making. <https://github.com/eleurent/highway-env>

# LogiCity: Advancing Neuro-Symbolic AI with **Abstract** Urban Simulation

**Concepts**

```

% Spatial
IsInInter (X)
LeftOf (X,Y)
IsClose (X,Y)
...

% Semantic
IsCar (X)
IsPedes (X)
IsAmbulance (X)
IsPolice (X)
...
    
```

**Rules**

```

Stop (X) :-IsAmbulance (Y) ,
           IsClose (X,Y)
Slow (X) :-IsPolice (Y) ,
          RightOf (X,Y)
    
```

**Agents**

% City1	% City2
A <sub>1</sub> : IsCar IsTruck	A <sub>1</sub> : IsCar
A <sub>2</sub> : IsCar IsTruck	A <sub>2</sub> : IsPedes IsOld
A <sub>3</sub> : IsCar IsAmbulance	A <sub>3</sub> : IsPedes IsOld
A <sub>4</sub> : IsCar	A <sub>4</sub> : IsCar IsPolice
...	...

## Grounded *Diverse* Different Cities





# LogiCity: Advancing Neuro-Symbolic AI with **Abstract** Urban Simulation

*Flexible* Use-configurable Rules!

## *Stop* Rule

```
Stop(X) :- CollidingClose(X,Y)  
Stop(X) :- IsAtInter(X), IsInInter(Y)  
Stop(X) :- IsAtInter(X), IsAtInter(Y),  
HigherPri(Y,X)
```

## *Slow* Rule

```
Slow(X) :- IsTiro(X), IsPedestrian(Y),  
NextTo(X,Y)  
Slow(X) :- IsPolice(X), IsYoung(Y),  
IsYoung(Z), NextTo(Y,Z), NextTo(X,Y)
```

## *Fast* Rule

```
Fast(X) :- IsReckless(X), IsCar(Y),  
IsAtInter(Y),  
HigherPri(X,Y)
```



# LogiCity Task: Sequential Decision Making



# LogiCity Task: Sequential Decision Making

Training Example



❄ Goal      ❄ Start      - - - Global Path

✓ Long-horizon (50~100 steps)

✓ Multi-Agent Interaction







Table 2: Empirical results of different methods in SPF task. TSR denotes trajectory success rate (most crucial) and DSR indicates decision success rate. <sup>†</sup> means Popper timed out. <sup>‡</sup> indicates conflict rules will be inducted for different actions. See our [website](#) and Appendix G for episode visualizations.

Supervision	Mode\Model	Easy			Medium			Hard			Expert		
		TSR	DSR	Score	TSR	DSR	Score	TSR	DSR	Score	TSR	DSR	Score
N/A	Oracle	1.00	1.00	8.51	1.00	1.00	8.45	1.00	1.00	9.63	1.00	1.00	4.33
	Random	0.07	0.00	0.00	0.06	0.00	0.00	0.04	0.01	0.00	0.05	0.06	0.00
Behavior Cloning	Popper [10]	<b>1.00</b>	1.00	8.51	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>	N/A <sup>†</sup>
	MaxSynth [35]	<b>1.00</b>	1.00	8.51	0.25	0.67	3.18	0.15	0.60	2.96	0.09	0.21	0.37
	HRI [11]	0.37	0.78	4.40	<b>0.48</b>	0.70	4.75	0.08	0.15	0.59	N/A <sup>‡</sup>	N/A <sup>‡</sup>	N/A <sup>‡</sup>
	NLM [7]	0.75	1.00	7.29	0.30	0.67	3.24	<b>0.24</b>	0.27	2.00	<b>0.22</b>	0.38	0.99
	GNN [37]	0.26	0.39	2.58	0.17	0.24	1.31	0.19	0.39	2.19	0.19	0.32	0.84
	MLP	0.61	0.63	4.80	0.20	0.19	1.22	0.12	0.13	0.81	0.10	0.19	0.25
Reinforcement Learning	NLM-DQN [7, 34]	<b>0.53</b>	0.96	5.93	<b>0.47</b>	0.67	4.40	<b>0.29</b>	0.40	2.69	<b>0.15</b>	0.35	0.62
	MB-shooting [36]	0.24	0.44	2.55	0.20	0.17	1.18	0.16	0.17	1.26	0.13	0.11	0.37
	DreamerV2 [38]	0.07	0.43	2.86	0.02	0.21	0.67	0.00	0.30	1.45	0.12	0.06	0.41
	DON [34]	0.35	0.89	4.80	0.42	0.59	3.72	0.09	0.12	0.63	0.07	0.24	0.37
	PPO [32]	0.33	0.36	2.83	0.09	0.25	0.88	0.02	0.38	1.57	0.12	0.08	0.38
	A2C [33]	0.10	0.16	1.00	0.06	0.29	1.07	0.00	0.14	0.46	0.12	0.09	0.34

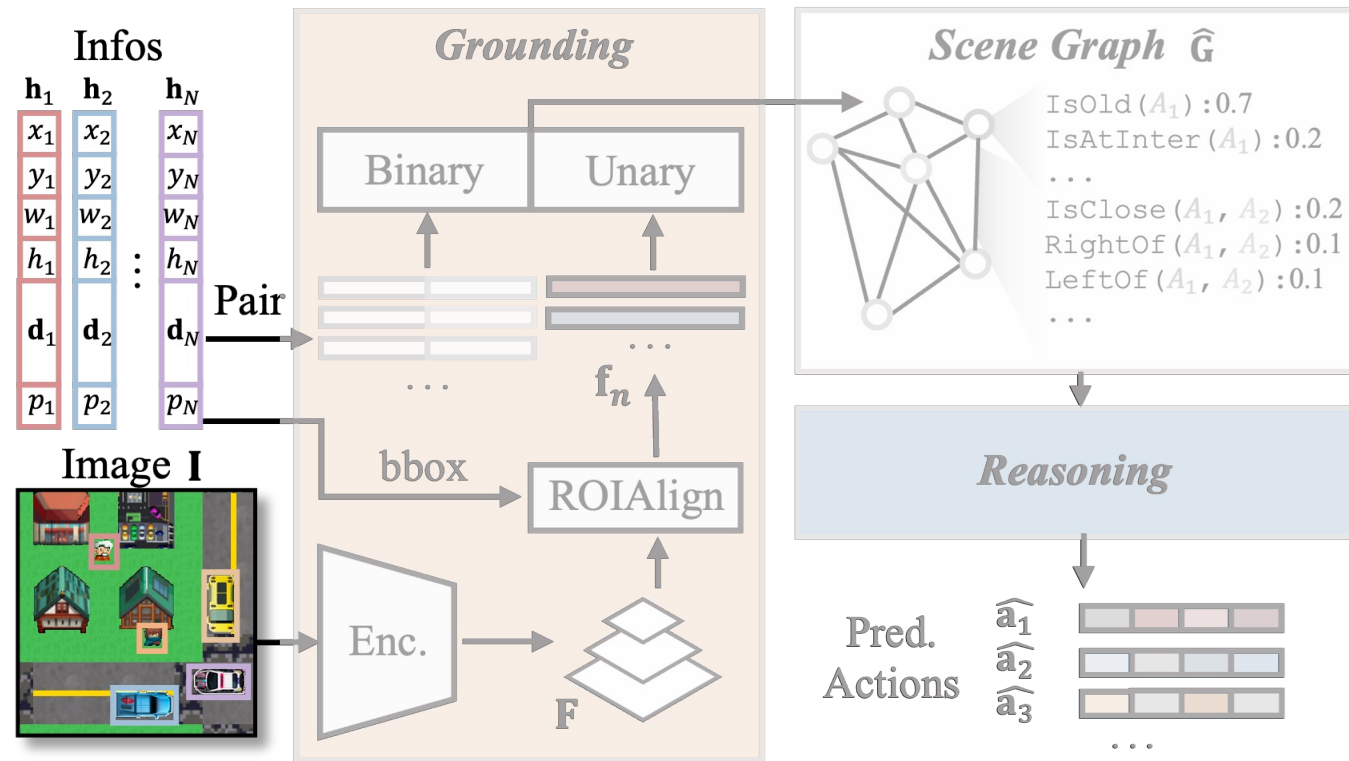




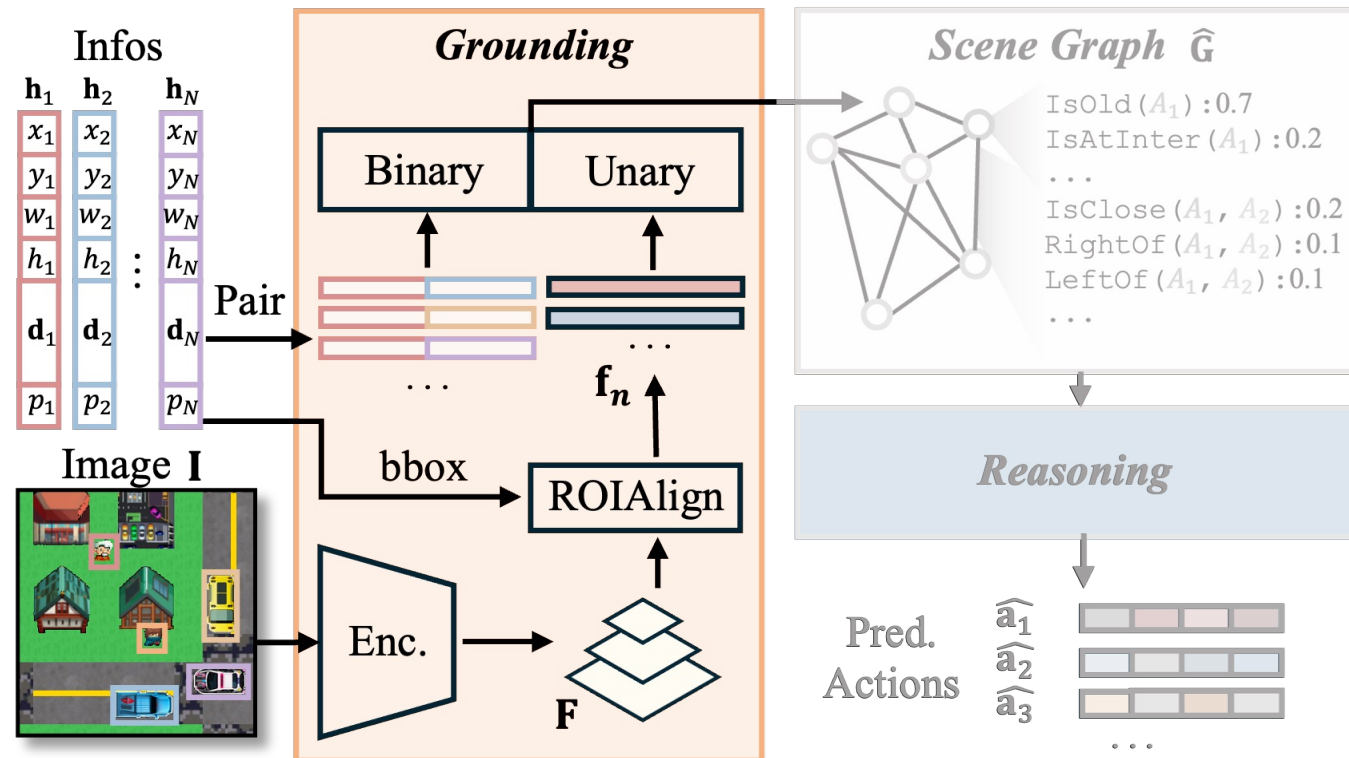
# LogiCity Task: Abstract Visual Reasoning



# LogiCity Task: Abstract Visual Reasoning

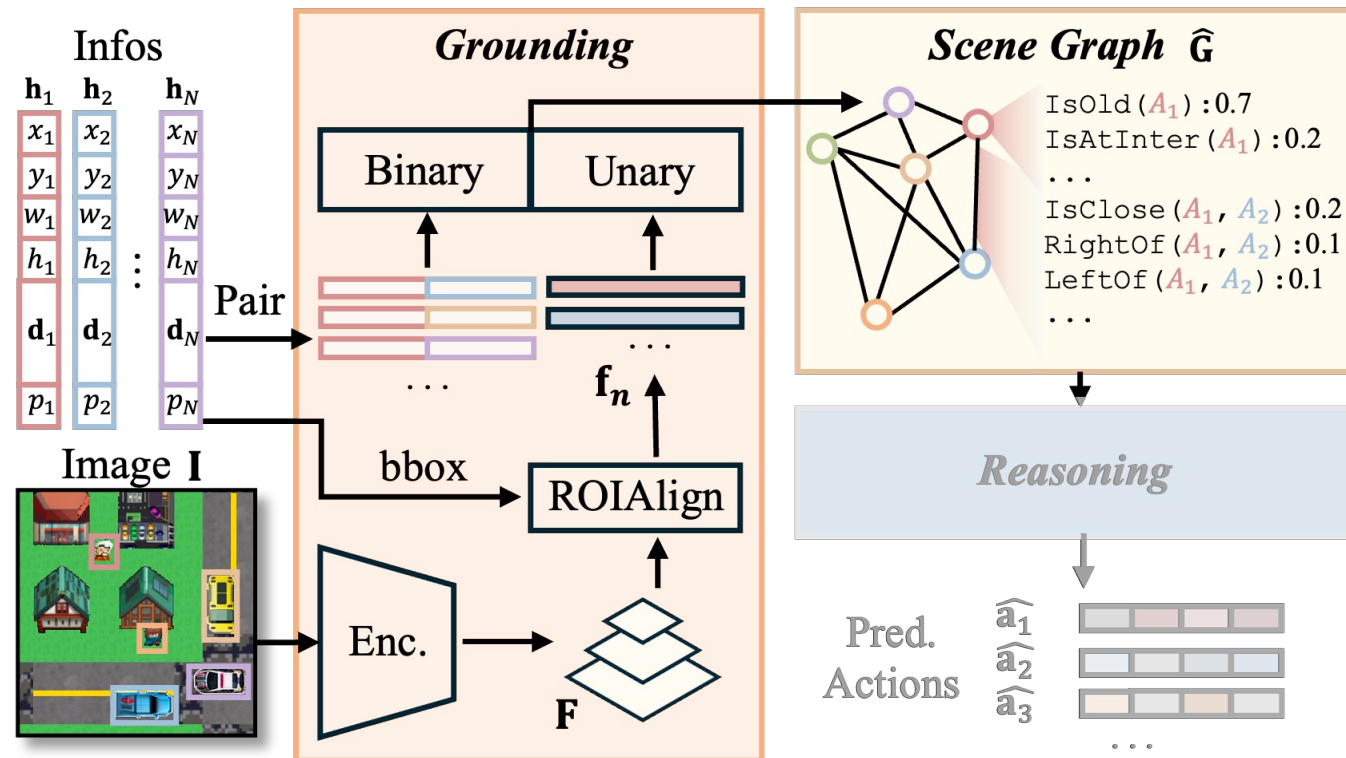


# LogiCity Task: Abstract Visual Reasoning

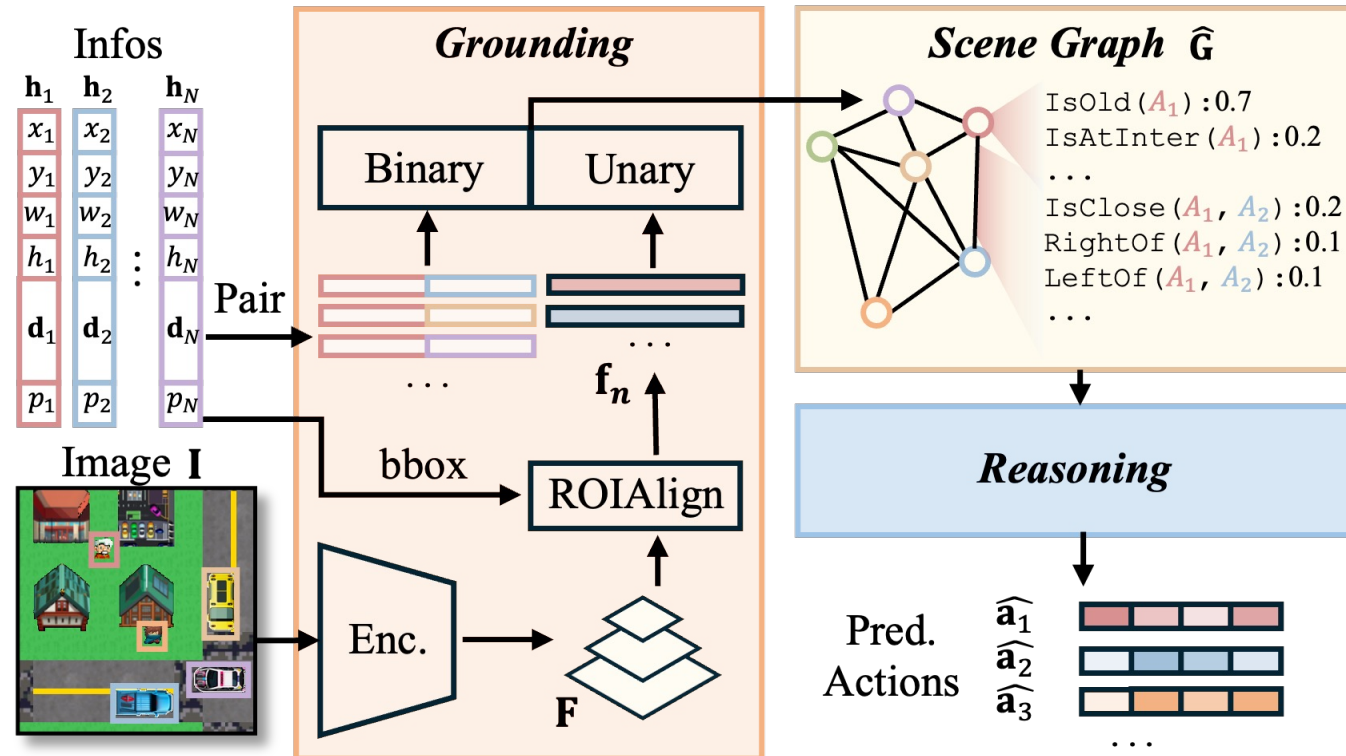




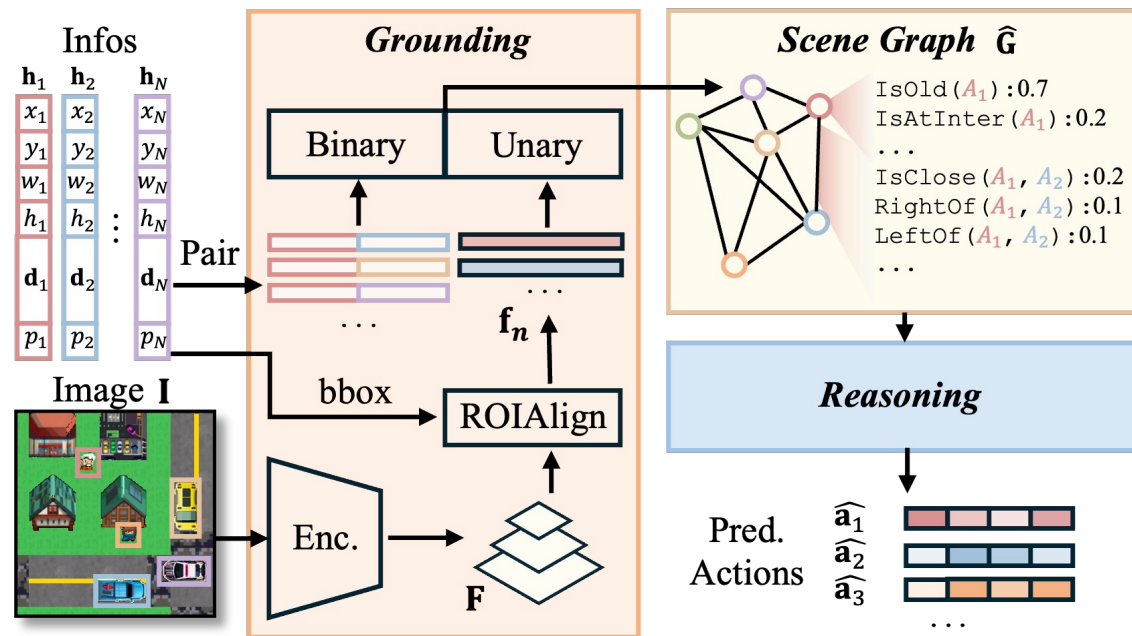
# LogiCity Task: Abstract Visual Reasoning



# LogiCity Task: Abstract Visual Reasoning



# LogiCity Task: Abstract Visual Reasoning



**Train**



**Test**





# LogiCity Task: Abstract Visual Reasoning



Grounded Clause

```
Normal (X) :-  
  IsReckless (X) ,  
  IsInInter (X)
```

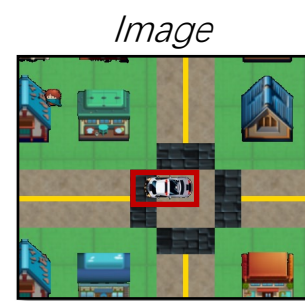


Predictions

GT: **Normal**

NLM: **Normal**

GNN: **Slow**



Grounded Clause

```
Slow (X) :-  
  Not (IsReckless (X)) ,  
  Not (IsAmbulance (X)) ,  
  IsInInter (X)
```



Predictions

GT: **Slow**

NLM: **Slow**

GNN: **Normal**



# LogiCity Task: Abstract Visual Reasoning

Image



Grounded Clause

Normal (X) :-  
 IsReckless (X) ,  
 IsInInter (X)

Predictions

GT: Normal

NLM: Normal

GNN: Slow



Stop (X) :-  
 IsAtInter (X) ,  
 IsInInter (Y) ,  
 Sees (X, Y)

GT: Stop

NLM: Stop

GNN: Normal



Normal (X) :-  
 Not (Stop (X))  
 IsReckless (X) ,  
 IsAtInter (Y) ,  
 Not (Sees (X, Y))

GT: Normal

NLM: Normal

GNN: Fast

Image



Grounded Clause

Slow (X) :-  
 Not (IsReckless (X)) ,  
 Not (IsAmbulance (X)) ,  
 IsInInter (X)

Predictions

GT: Slow

NLM: Slow

GNN: Normal



Fast (X) :-  
 Not (Stop (X))  
 IsReckless (X) ,  
 IsAtInter (Y) ,  
 Sees (X, Y)

GT: Fast

NLM: Fast

GNN: Normal



Slow (X) :-  
 IsPolice (X) ,  
 IsYoung (Y) ,  
 IsYoung (Z) ,  
 NextTo (Y, Z)  
 Sees (X, Y)

GT: Slow

NLM: Slow

GNN: Normal



# LogiCity: Advancing Neuro-Symbolic AI with Abstract Urban Simulation

Bowen Li<sup>1</sup> Zhaoyu Li<sup>2</sup> Qiwei Du<sup>3</sup> Jinqi Luo<sup>4</sup> Wenshan Wang<sup>1</sup> Yaqi Xie<sup>1</sup> Simon Stepputtis<sup>1</sup>  
Chen Wang<sup>3</sup> Katia Sycara<sup>1</sup> Pradeep Ravikumar<sup>1</sup> Alexander Gray<sup>5</sup> Xujie Si<sup>2</sup> Sebastian Scherer<sup>1,\*</sup>

<sup>1</sup>Carnegie Mellon University, <sup>2</sup>University of Toronto, <sup>3</sup>University at Buffalo  
<sup>4</sup>University of Pennsylvania, <sup>5</sup>Centaur AI Institute

NeurIPS 2024 D&B track (w/ audio)

