Adversarial Locomotion and Motion Imitation for Humanoid Policy Learning

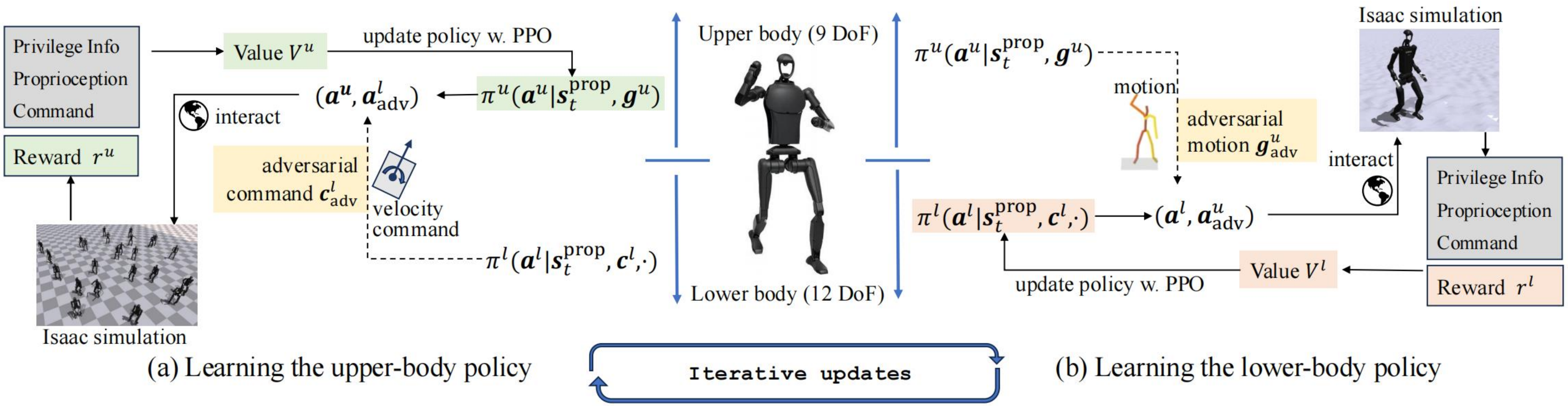
Jiyuan Shi[1*], Xinzhe Liu[1,2*], Dewei Wang[1,3*], Ouyang Lu[1,4], Sören Schwertfeger[2], Fuchun Sun[5], Chi Zhang[1], Chenjia Bai[1†], Xuelong Li[1†]

[1]Institute of Artificial Intelligence (TeleAI), China Telecom [2]ShanghaiTech University
[3]University of Science and Technology of China [4]Northwestern Polytechnical University
[5]Tsinghua University

[*]Equal Contribution [†]Corresponding Author

We propose a novel **adversarial** training framework (**ALMI**) for **robust locomotion** and **precise motion imitation** for humanoid robots.

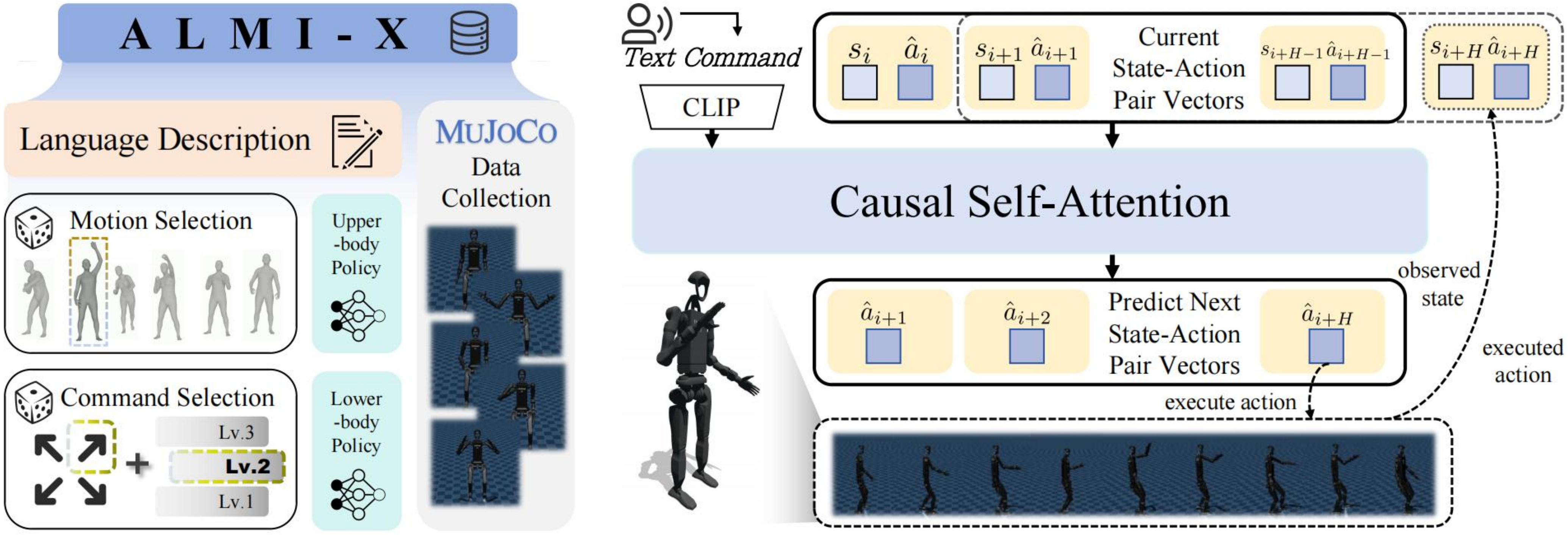
- **Adversarial** training framework that **iteratively** updates **lower** and **upper-body policies**.
- **Dual curriculum mechanism** for **stable training** and **sampling adversarial actions**.



(a) Learning the upper-body policy (b) Learning the lower-body policy

We release a **large-scale whole-body motion control dataset** named **ALMI-X** featuring **high-quality episodic trajectories**.

- Contain 1989 motions combined with 41 commands, resulting **81,549 trajectories** totally.
- Give a try to train a **transformer-based whole-body foundation model** using ALMI-X.



(a) ALMI-X Collection (b) Transformer-based Foundation Policy

Experiment Results

We conduct extensive experiments in both the simulation and the real world.

Table 2: Simulated evaluation of ALMI, ALMI (whole body) and Exbody on CMU dataset.

Method	Metrics							
	$E_{\text{vel}} \downarrow$	$E_{\text{ang}} \downarrow$	$E_{\text{jpe}}^{\text{upper}} \downarrow$	$E_{\text{kpe}}^{\text{upper}} \downarrow$	$E_{\text{action}}^{\text{upper}} \downarrow$	$E_{\text{action}}^{\text{lower}} \downarrow$	$E_{\text{g}} \downarrow$	Survival $\uparrow$
Easy								
ALMI	**0.1135**	**0.2647**	**0.1931**	**0.0460**	**0.0462**	**0.0170**	**0.6919**	**1.0000**
ALMI(whole)	0.1386	0.5433	0.5756	0.0704	0.0800	3.0356	0.9675	0.9991
Exbody	0.2383	0.4056	0.3559	0.0995	1.7813	1.8152	0.9693	0.8912
Medium								
ALMI	**0.2192**	**0.3520**	**0.2007**	**0.0450**	**0.0598**	**0.0172**	**0.7604**	**0.9852**
ALMI(whole)	0.2380	0.5563	0.6734	0.0637	0.0409	2.9225	1.0750	0.9763
Exbody	0.3063	0.5087	0.3658	0.1233	1.7683	1.8019	1.0166	0.8845
Hard								
ALMI	**0.2202**	**0.4812**	**0.2116**	**0.0458**	**0.0600**	**0.0175**	**0.8551**	**0.9723**
ALMI(whole)	0.3178	0.7224	0.7022	0.0635	0.0519	2.9317	1.1656	0.9491
Exbody	0.4838	0.5753	0.3758	0.1269	1.7352	1.7689	1.0243	0.8778

Table 3: Ablation studies of adversarial training technique and arm curriculum in ALMI.

Method	Metrics							
	$E_{\text{vel}} \downarrow$	$E_{\text{ang}} \downarrow$	$E_{\text{jpe}}^{\text{upper}} \downarrow$	$E_{\text{kpe}}^{upper} \downarrow$	$E_{\text{action}}^{\text{upper}} \downarrow$	$E_{\text{action}}^{\text{lower}} \downarrow$	$E_{\text{g}} \downarrow$	Survice $\uparrow$
Easy								
lower-3 + upper-2	**0.1135**	**0.2647**	0.1931	**0.0460**	**0.0462**	**0.0170**	**0.6919**	**1.0000**
lower-2 + upper-2	0.1164	0.2669	0.1955	0.0452	0.0475	**0.0171**	0.7121	**1.0000**
lower-1 + upper-2	0.1271	0.2738	0.1928	0.0526	0.0642	**0.0171**	0.7052	**1.0000**
w/o arm curriculum	0.1411	0.2726	**0.1924**	0.0504	0.0618	**0.0172**	0.7472	0.9995
Medium								
lower-3 + upper-2	**0.2192**	**0.3520**	**0.2007**	**0.0450**	**0.0598**	**0.0172**	**0.7604**	**0.9852**
lower-2 + upper-2	0.2213	0.3571	0.2032	0.0458	0.0607	**0.0172**	0.7748	0.9772
lower-1 + upper-2	0.2262	0.3872	0.2173	0.0492	0.0604	0.0175	0.7730	0.9273
w/o arm curriculum	0.2571	0.4348	0.2068	0.0476	0.0601	0.0173	1.0587	0.9652
Hard								
lower-3 + upper-2	**0.2202**	**0.4812**	**0.2116**	**0.0458**	**0.0600**	**0.0175**	**0.8551**	**0.9723**
lower-2 + upper-2	0.2892	0.5395	0.2231	0.0482	0.0645	0.0178	0.9479	0.9233
lower-1 + upper-2	0.2566	0.5172	0.2451	0.0537	0.0777	0.0179	0.9462	0.8743
w/o arm curriculum	0.3658	0.6398	0.2394	0.0461	0.0726	0.0180	1.2042	0.8480

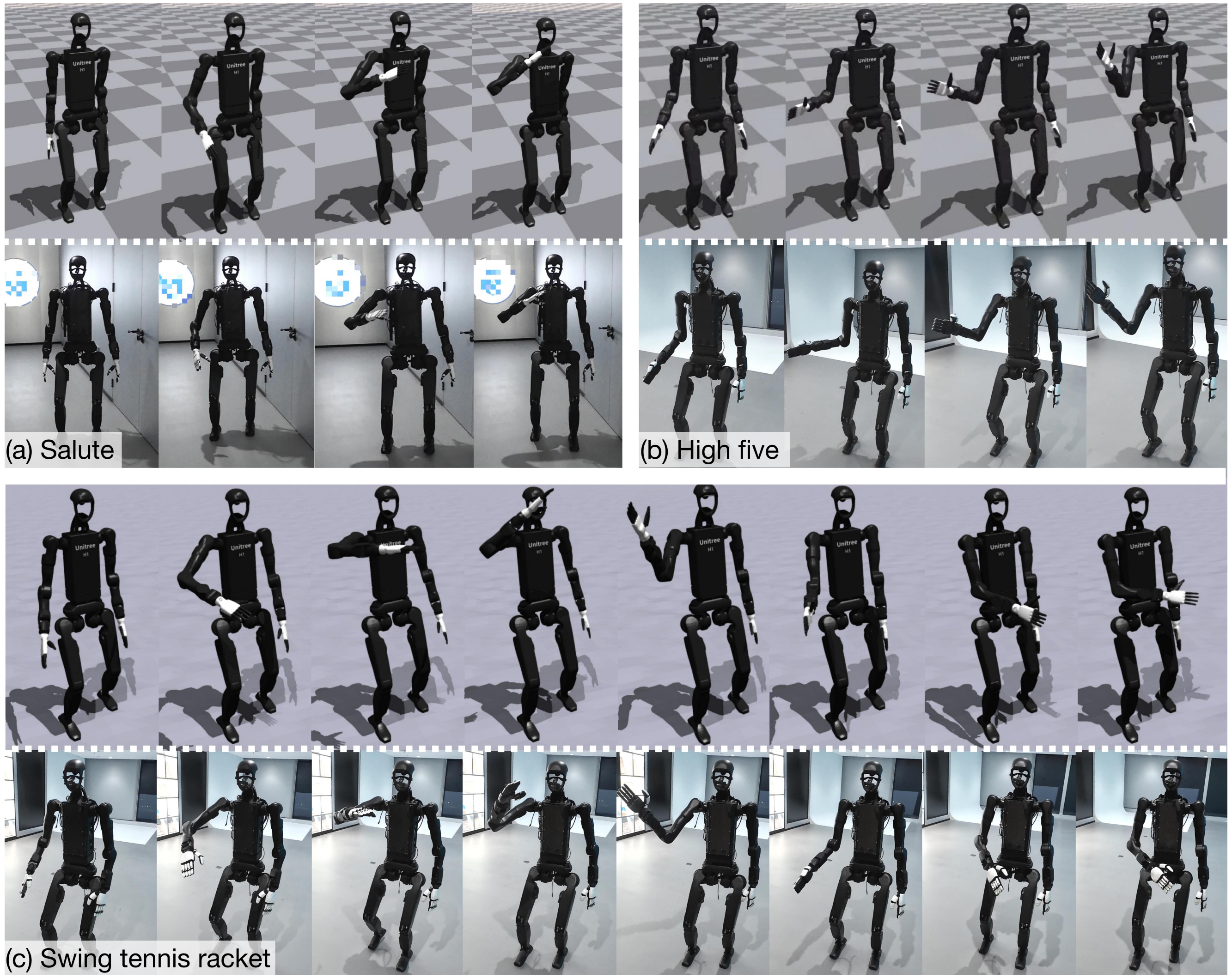
Table 11: Average velocity of Foundation models with different text commands.

Text commands w.r.t. lower body	Average Velocity		
	$\bar{v}_x$	$\bar{v}_y$	$\bar{\omega}_{\text{yaw}}$
CL-20sl			
forward fast	0.48	0.18	0.08
left fast	0.07	0.28	0.00
backward to left slowly	-0.19	0.21	0.00
turn left slowly	0.02	0.02	0.08
go right fast	-0.11	-0.37	-0.07
keep standing	0.00	0.00	0.01
CL-400sl			
forward fast	0.87	-0.05	-0.10
left fast	0.53	0.42	-0.07
backward to left slowly	0.35	0.30	-0.01
turn left slowly	0.53	0.16	0.2
go right fast	0.47	-0.50	-0.10
keep standing	0.22	0.08	0.00

Table 12: Survival duration (SD) and success rates of upper-body(SR_{up})/lower-body(SR_{low}) with different text commands.

Text commands	Metrics		
	SR_{low}	SR_{up}	SD
CL-20sl			
go forward slowly and wave left.	1.00	0.20	8.0
go backward moderately and wave right.	1.00	0.20	8.0
go right fast and wave both.	1.00	0.40	8.0
CL-400sl			
go forward slowly and wave left.	1.00	1.00	2.14
go backward moderately and wave right.	1.00	1.00	2.54
go right fast and wave both.	0.00	1.00	3.57
OL			
go forward slowly and wave left.	0.00	1.00	0.31
go backward moderately and wave right.	0.00	1.00	0.29
go right fast and wave both.	0.00	1.00	0.32

We conduct extensive experiments in both the simulation and the real world.




Kitchen Manipulation Tasks
TeleAI
具身智能团队
Catch apple
1X