End-to-End Language-Guided RL for Legged Robots

Raül Ojeda Gandia

ARF - 24/25

May 2025

Introduction & Motivation

Towards Generalist Robotics

The Intelligence Gap

- Large Language Models demonstrate near-human intelligence in software
- Physical intelligence in robotics remains limited and task-specific

Why Language-Driven Robot Control?

- Language is humanity's interface for complex reasoning
- Enables robots to leverage semantic understanding from LLMs

Social & Environmental Impact

- Social: Generalist robots could assist in diverse daily tasks
- Environmental: Single adaptable robots replace multiple specialized ones

Related Work

Traditional Approaches

- Rule-based systems utilizing predefined motion primitives
- Typically limited to narrow command sets and specific contexts

Recent Advances

- **IsaacLab Environment**: Provides a scalable simulation platform for velocity tracking and control policy training – forms the foundation of our work
- **Kumar et al. (2023)** – *Words into Action: Learning Diverse Humanoid Robot Behaviors using Language-Guided Iterative Motion Refinement*
- **Sun et al. (2024)** – *Leveraging Large Language Models for Comprehensive Locomotion Control in Humanoid Robots*

Task Description

Robot

- Humanoid robot Unitree G1, 1.20 meters tall, 35 kg, 37 DOF

Data Acquisition Process

- Generated diverse natural language movement commands using Claude.ai
- Each command paired with target velocities (linear_x, linear_y, angular_z)
- Pre-computed transformer embeddings stored as .npy files
- Progressive dataset expansion to address learning imbalances

Dataset Statistics	Value
Total commands	922
English	598 (65%)
Spanish	142 (15%)
German	76 (8%)
French	68 (7%)
Other	38 (4%)

Movement Distribution	Count
Forward commands	412 (45%)
Backward commands	125 (14%)
Turn commands	198 (22%)
Stop commands	89 (10%)
Lateral commands	62 (7%)
Combined movements	36 (4%)

Feature Extraction & Input/Output

Language Embedding

- Model: `all-MiniLM-L6-v2` (22M params)
- Optimized for semantic similarity

Observation Space (691D)

- **Motion:** Linear/Angular Velocity $v, \omega \in \mathbb{R}^3$, Gravity Vector $g \in \mathbb{R}^3$
- **Joints:** Positions/Velocities $\theta, \dot{\theta} \in \mathbb{R}^{37}$
- **Context:** Language Embedding $e \in \mathbb{R}^{384}$, Height Scan $h \in \mathbb{R}^{187}$, Previous Action $a_{t-1} \in \mathbb{R}^{37}$

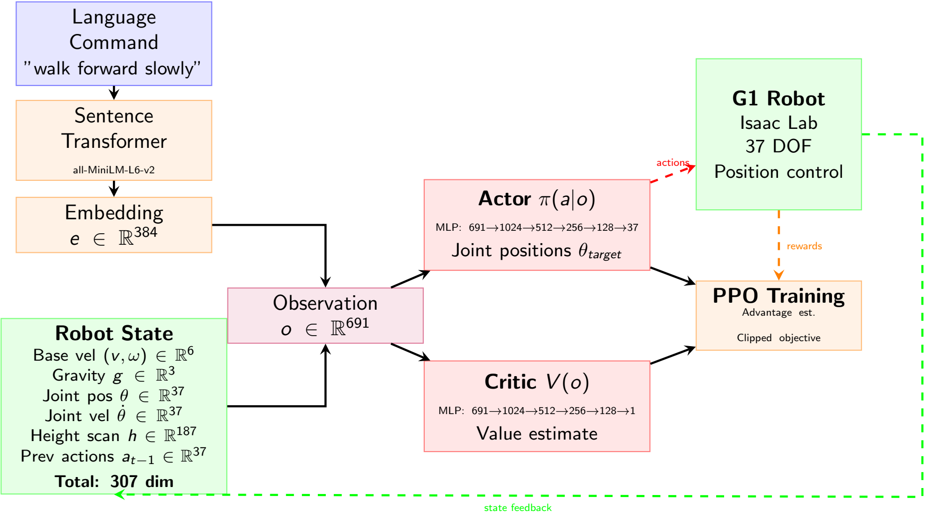
Output

- Joint Position Targets $\theta_{target} \in \mathbb{R}^{37}$

Reward Components

- **Task:** Language tracking (2.5), Success bonus (1.5)
- **Regularization:** Torques, accelerations, smoothness
- **Gait Quality:** Air time, orientation, joint stability

Model Architecture



Experimental Design

Training Infrastructure

- Environment: Isaac Lab with 4096 parallel environments
- Hardware: Laptop 4070 GPU, 32 GB RAM, achieving 38,000+ steps/second
- Total training time: $\sim$20 hours across multiple sessions

Progressive Training Curriculum

1. **Phase 1 (0-3k iterations)**: Basic movement
   - 30 simple slow-movement commands only
   - Robot learned stability but minimal movement
2. **Phase 2 (3k-10k iterations)**: Expanded command set
   - 90 commands: slow + fast movements, all directions
   - Added anti-standing penalty to encourage motion
3. **Phase 3 (10k-17k iterations)**: Multilingual generalization
   - 600 commands with linguistic variations
   - Multiple languages (EN, ES, DE, FR, ZH)
4. **Phase 4 (17k-25k iterations)**: Forward motion correction
   - Added 300+ forward movement commands
   - Addressed directional learning imbalance
   - Final dataset: 922 total commands

Experimental Design: Test Set Evaluation

Test Set Composition (120 commands)

- **Close variations**: 25
- **Novel concepts**: 19
- **Novel verbs**: 15
- **Multilingual**: 10
- **Slang/Informal**: 10
- **Typos**: 9
- **Formal/Technical**: 10
- **Uncertain**: 8
- **Abbreviations**: 7
- **Emphasis**: 7

Movement Distribution

- Forward: 45 commands
- Backward: 12
- Left/Right turns: 30
- Stop: 15
- Lateral: 8
- Combined: 10

Results: Performance by Category

Performance by Test Category

Test Category	Success	Failure
Close variations (25)	19	6
Novel concepts (19)	13	6
Novel verbs (15)	8	7
Multilingual (10)	4	6
Slang/Colloquial (10)	4	6
Typo variations (9)	3	6
Formal/Technical (10)	6	4
Uncertain/Hedged (8)	4	4
Emphasis/Urgency (7)	4	3
Abbreviations (7)	4	3

Performance by Movement Type

Movement	Success	Failure
Forward (45)	13	32
Left turn (15)	10	5
Right turn (15)	7	8
Stop (15)	10	5
Backward (12)	9	3
Combined (10)	8	2
Lateral (8)	5	3

Overall Test Set Performance

Total Commands: 120 Success: **69/120 (58%)**

Discussion: Generalization Analysis

Successful Generalizations

- ✓ "robot march forward" → Correct forward motion
- ✓ "trun lft" → Handled typos correctly
- ✓ "bkwd" → Understood abbreviation
- ✓ "gira a destra per favore" → Cross-lingual transfer
- ✓ "go back home" → Novel concept

Failed Generalizations

- ✗ "navigate to my position" → No movement
- ✗ "if you dont mind go right" → Sliding right instead of rotating right
- ✗ "scoot forward" → Novel verb
- ✗ "patrol forward and back" → Only forward

Conclusions

Key Achievements

- Achieved 58% success rate on diverse test set (120 novel commands)
- Demonstrated emergent capabilities

Technical Insights

- Direct language-to-action mapping works without symbolic intermediates
- Progressive curriculum essential for balanced learning

Towards Generalist Robotics

Natural language provides a scalable interface for robot control,
enabling semantic generalization beyond explicit training data

Future Work

Immediate Improvements

- Address forward motion bias through improved reward shaping
- Incorporate larger language models for better semantic understanding

Physical Robot Deployment

- Transfer learned policies to real G1 hardware

Advanced Capabilities

- Vision-language grounding: "walk to the red object"
- Integration with LLM planners for complex reasoning

Long-term Vision

From single commands to conversational robot control
enabling truly generalist physical AI systems

Thank you!

Questions?