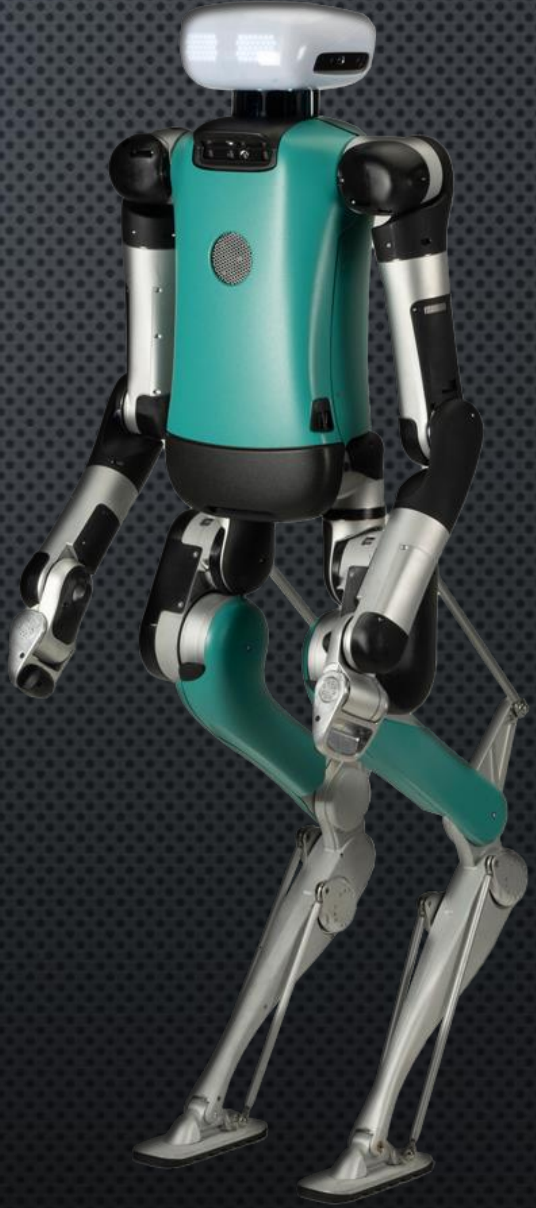
Real-World Humanoid Locomotion with Reinforcement Learning

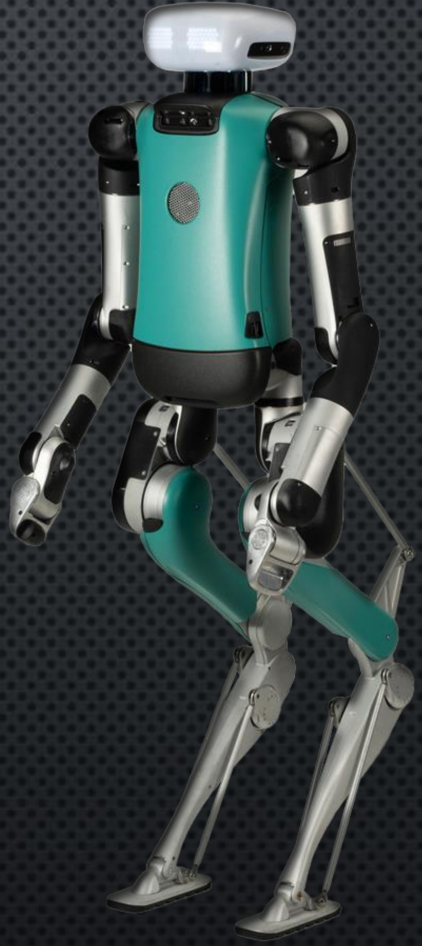
Ilija Radosavovic

Group 28



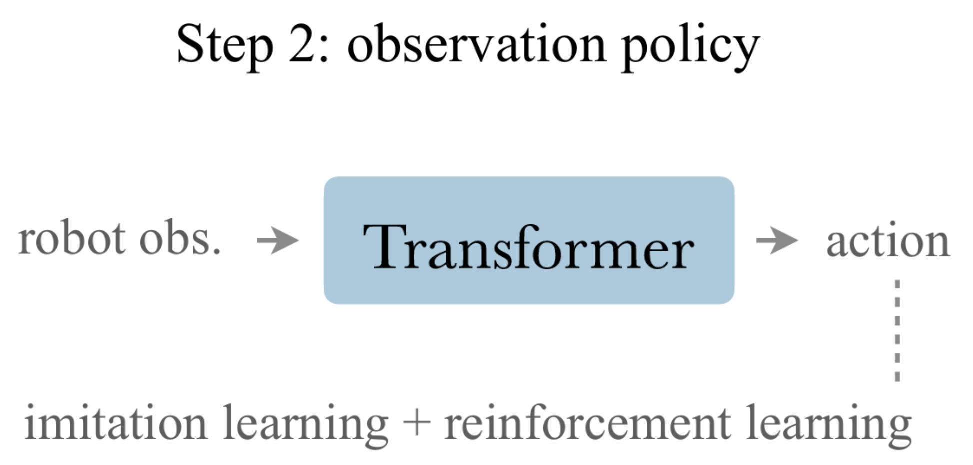
PLAN

I. Summary of the main idea/contribution
II. Key aspects
III. The article
IV. Who is citing the article ?
V. Analysis of pros and cons
VI. Possible exam questions



SUMMARY OF THE MAIN IDEA

- **Reinforcement learning approach** for real world humanoid locomotion
- Use of a **causal transformer** to predict actions based only on proprioceptive sensors
  - Hypothesis: Behavior adaptation based on observation-action history

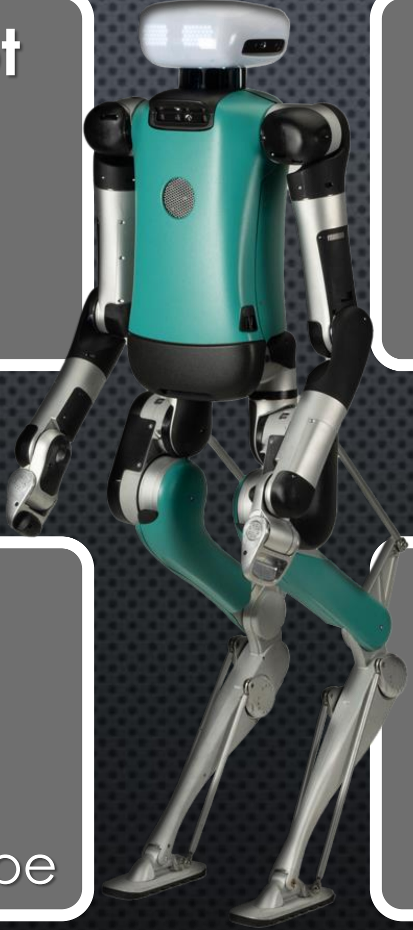


- **Contribution**: **Hope** for **future work** and **research**
  - Autonomous navigation in complex and dynamic terrains using a simple and general learning-based controller
  - Encourage future work

KEY ASPECTS

Biped humanoid robot
Agility Robotics
1,6m tall
45kg
30 DOF

Control
Model-free reinforcement learning
Causal transformer

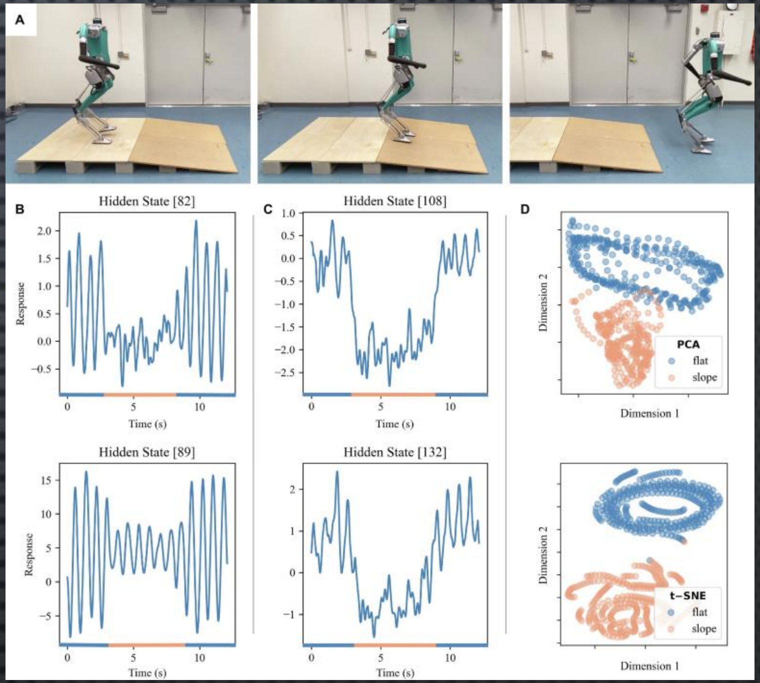


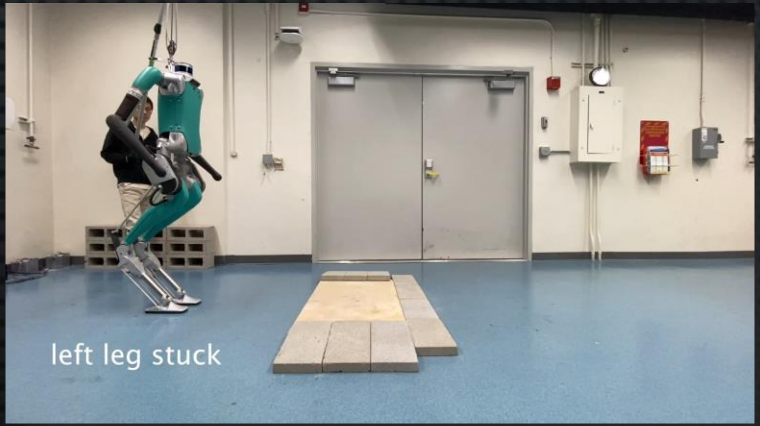
Gait type
Normal walking
Fast walking
Omnidirectional walking
Adapting on the terrain/slope

Sensors
Input: History of observations and actions
Joint encoders
IMU

THE ARTICLE - RESULTS

- Stable walking across many conditions
- Emergent behavior
  - Gait changes
  - Foot recovery
  - Arm-swing
- Robustness to external disturbance



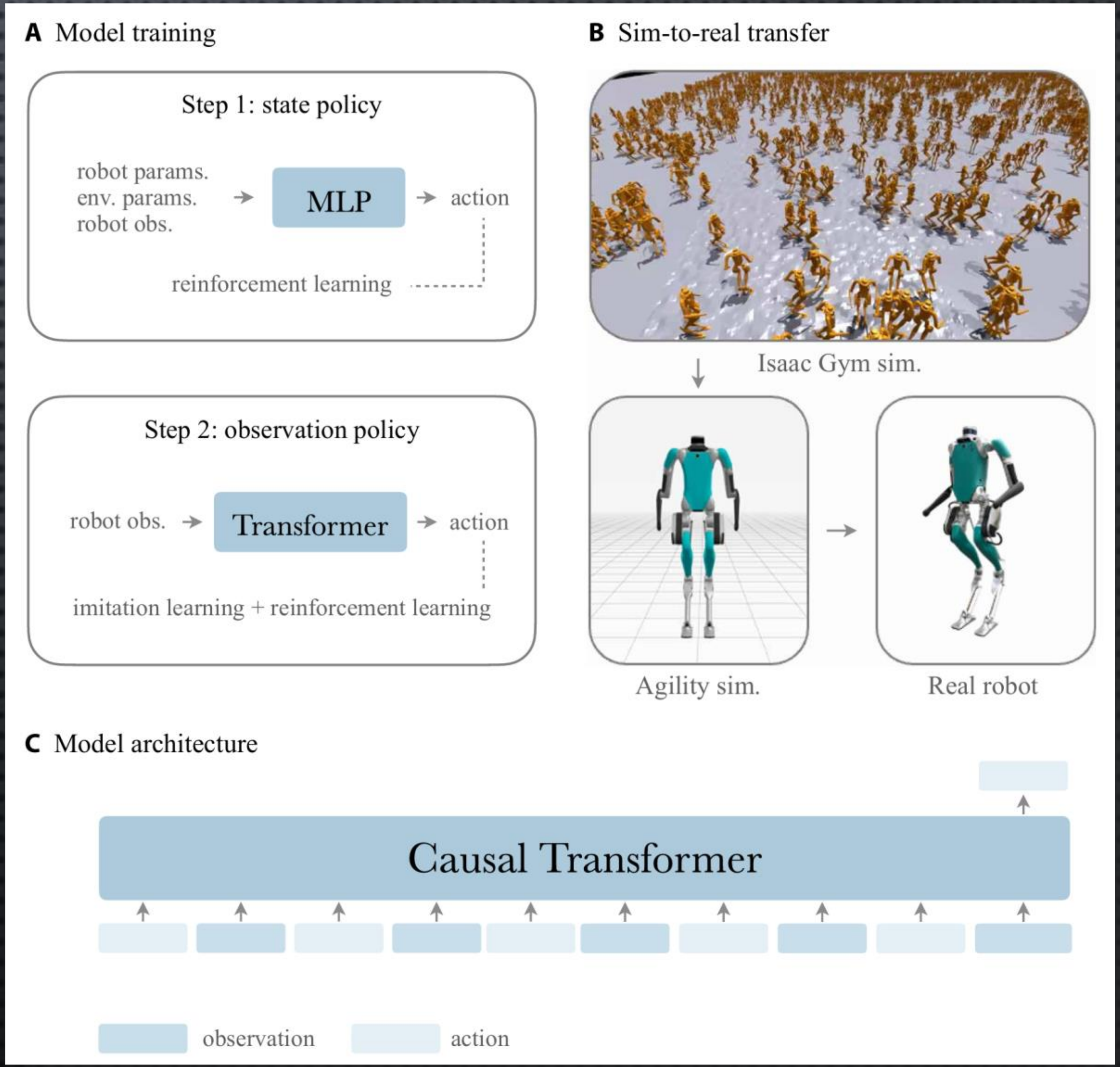


THE ARTICLE - CONTROLLER

- Causal transformer
- Training
  1. Isaac Gym simulator
     - Teacher policy
     - Student policy
  2. Agility Simulator
  3. Real robot

- 50Hz neural network policy
- Command to 1kHz PD controller

- Promising results for the future



WHO IS CITING THE ARTICLE ?

- 2024: citations in 7 different articles
  - *A survey on legged robots: Advances, technologies and applications*, Wu, Z., Zheng, K., Ding, Z., Gao, H.
  - *Realizing full-body control of humanoid robots*, Li, G., Gomez, R.
  - *A Robust Strategy for UAV Autonomous Landing on a Moving Platform under Partial Observability*, Aikins, G., Jagtap, S., Nguyen, K.-D.
  - *Robust humanoid robot vehicle ingress with a finite state machine integrated with deep reinforcement learning*, Wang, C., Chen, X., Yu, Z., (...), Chen, K., Gergondet, P.
  - *A Survey of Optimization-Based Task and Motion Planning: From Classical to Learning Approaches*, Zhao, Z., Cheng, S., Ding, Y. Xu, D., Zhao, Y.
  - *Load-Carrying Assistance of Articulated Legged Robots Based on Hydrostatic Support*, Fan, W., Dai, Z., Li, W., Liu, T.
  - *Not only Rewards but Also Constraints: Applications on Legged Robot Locomotion*, Kim, Y., Oh, H., Lee, J., (...), Youm, D., Hwangbo, J.

WHO IS CITING THE ARTICLE ?

- 2024: citations in 7 different articles
  - *A survey on legged robots: Advances, technologies and applications*, Wu, Z., Zheng, K., Ding, Z., Gao, H.
  - *Realizing full-body control of humanoid robots*, Li, G., Gomez, R.
  - *A Robust Strategy for UAV Autonomous Landing on a Moving Platform under Partial Observability*, Aikins, G., Jagtap, S., Nguyen, K.-D.
  - *Robust humanoid robot vehicle ingress with a finite state machine integrated with deep reinforcement learning*, Wang, C., Chen, X., Yu, Z., (...), Chen, K., Gergondet, P.
  - *A Survey of Optimization-Based Task and Motion Planning: From Classical to Learning Approaches*, Zhao, Z., Cheng, S., Ding, Y. Xu, D., Zhao, Y.
  - *Load-Carrying Assistance of Articulated Legged Robots Based on Hydrostatic Support*, Fan, W., Dai, Z., Li, W., Liu, T.
  - *Not only Rewards but Also Constraints: Applications on Legged Robot Locomotion*, Kim, Y., Oh, H., Lee, J., (...), Youm, D., Hwangbo, J.

Success of reinforcement learning in legged robot

Isaac Gym simulator for successful sim-to-real transfer

Simple example of learning-based approach

Unsuitable for advanced applications : vehicle ingress

- Only trained on the locomotion task
- Reward function with the same parameters

Low energy efficiency and weak loading capacity

- Unable to achieve large-scale applications (manufacturing and services)
- Compared to wheeled robots

Engineering of the controller

- Complex reward engineering process
- Time-consuming
- Nontrivial
- Dependence on engineer’s experience/intuition

ANALYSIS OF PROS AND CONS

- **Adapt** its behavior **dynamically** based on **context**
  - Gait changes
  - Rapid adaptation to sudden obstacles
  - Unlike gradient based methods, requiring fine-tuning on task-specific data samples, transformers can learn in context (ability to handle diverse inputs)
- **Good results**, **outperforms** other controllers in many situations
  - Reliable outdoor walking without falls
  - Robust to external disturbance
  - Different terrains
  - Varying payloads
  - High speed locomotion
  - Arm-swing motion for energy reduction

Unsuitable for advanced applications : vehicle ingress

Low energy efficiency and weak loading capacity

Engineering of the controller

- Needs to be **completed** by other **sensors** like cameras to avoid bumps and foot-trapping
- **Not perfectly symmetrical** policy
  - Different trajectories for 2 motors
  - Better at lateral movements to the left
- Policy **not perfect** at tracking the **commanded velocity**
- Robot **could fall** under excessive external disturbances

POSSIBLE EXAM QUESTIONS

What is the main advantage of using reinforcement learning ? Examples…

The generation of emergent behaviors / dynamic adaptability to the context
(gait changes, foot recovery, arm-swing)

What is the main reason for the emergence of the arm-swing movement in RL controller ?

Energy minimization as part of the reward function during the design of the controller