Humanoid Robotics

Locomotion 4:
Review on Locomotion

Maren Bennewitz, Shahram Khorshidi

Goals of this Lecture

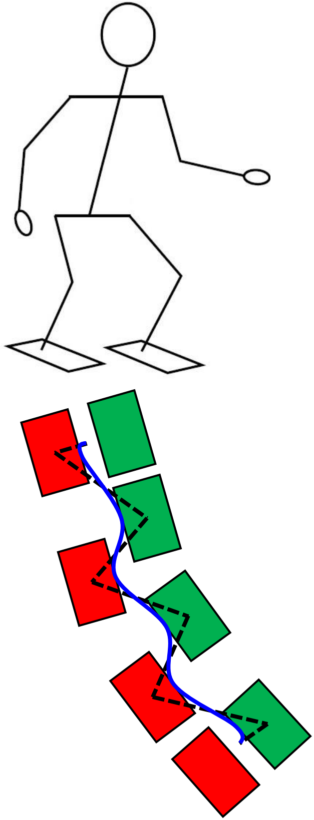
- Review the main components of the **legged-locomotion** control pipeline
- Revisit the role of **optimization-based methods** in trajectory generation, MPC, and control
- A brief introduction on **learning-based approaches** for locomotion planning and control.

Optimization-Based Legged Locomotion

Legged Locomotion

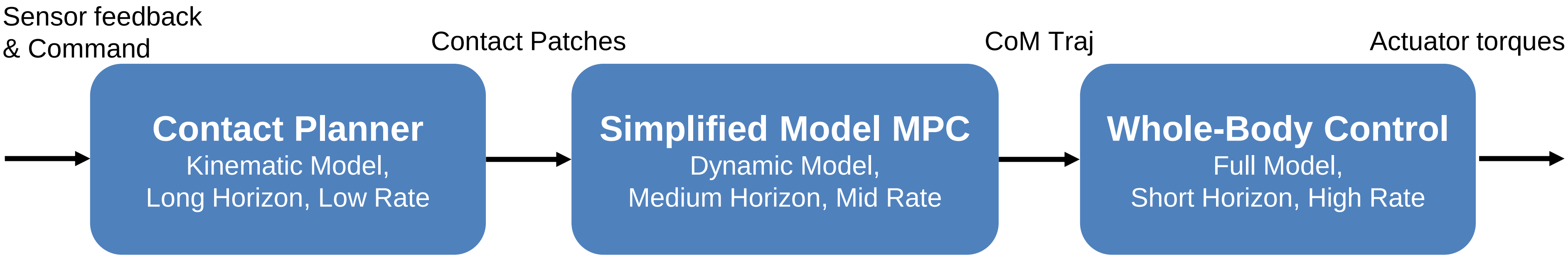
- **Decomposition** approach

Footsteps → Desired ZMP Trajectory → CoM Trajectory → Joint Commands →



Legged Locomotion

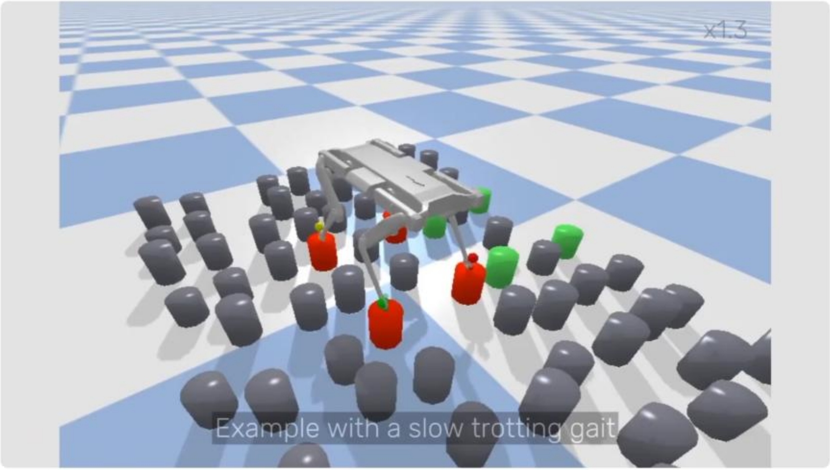
- **Breaking** the locomotion problem into smaller pieces, in order to be computationally trackable.



Contact Planning

- **Discrete** and **combinatorial** problem: decide which contacts to use, where to place them, and in what sequence.
- Contact choices must satisfy **kinematic** reachability, **collision** avoidance, **terrain** geometry, and stability constraints.
- **Sampling-based methods** are useful for irregular terrain or large, continuous contact spaces.

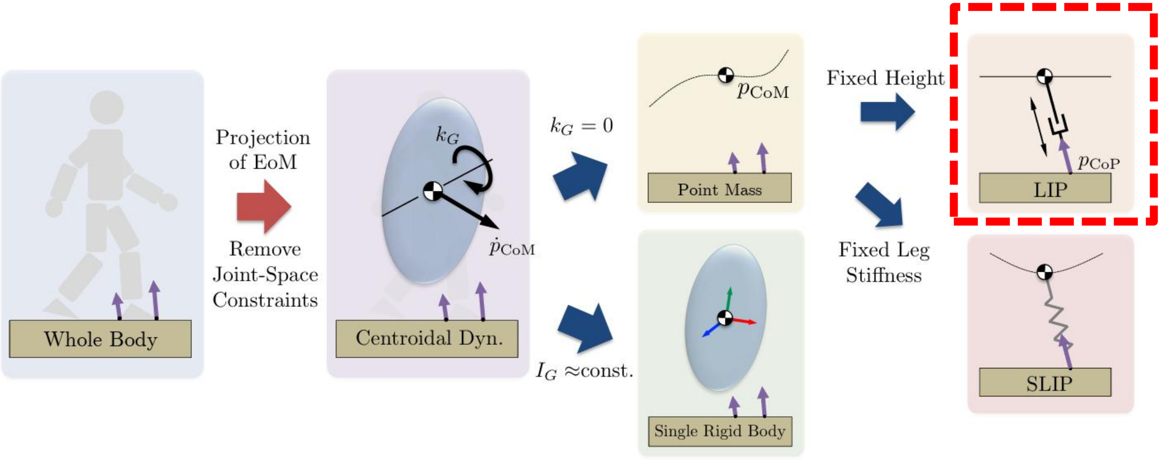
Contact Planning



Victor Dh´Edin, et al. “Diffusion-based learning of contact plans for agile locomotion,” Humanoids, 2024.

Motion Planning - Modeling Choices

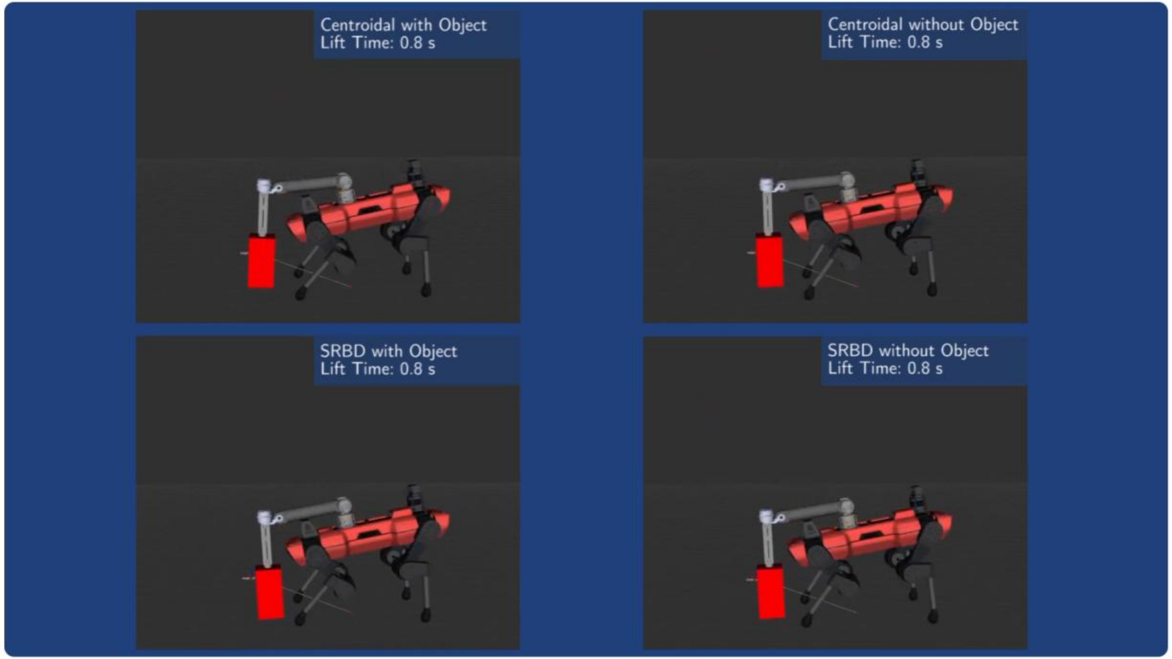
- Simplified models retain the **dominant** locomotion dynamics while removing joint-level complexity.
- The main trade-off is **model fidelity** versus **computational efficiency** and optimization tractability.



Wensing, et al. “Optimization-based control for dynamic legged robots,” IEEE TRO, 2024.

Motion Planning - Modeling Choices

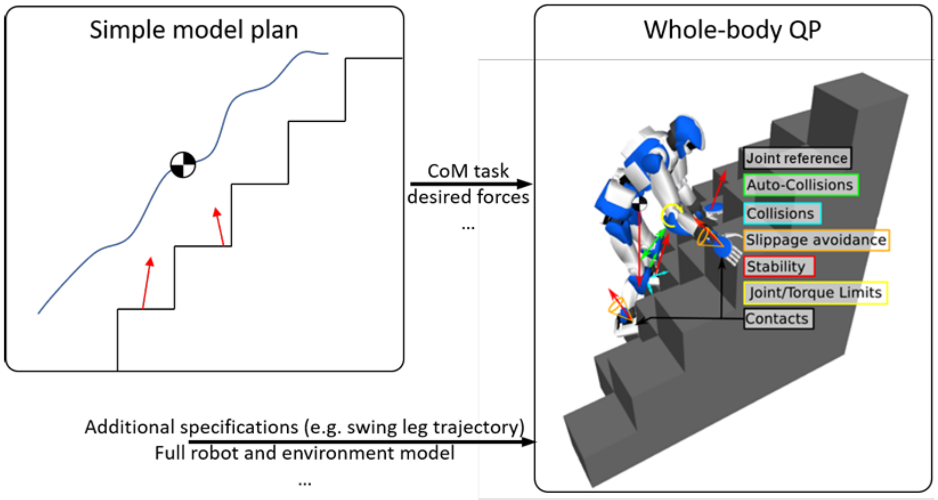
- **Loco-manipulation** MPC with centroidal dynamics



Sleiman, et al. “A Unified MPC Framework for Whole-Body Dynamic Locomotion and Manipulation,” IEEE ICRA, 2021.

Whole Body Control

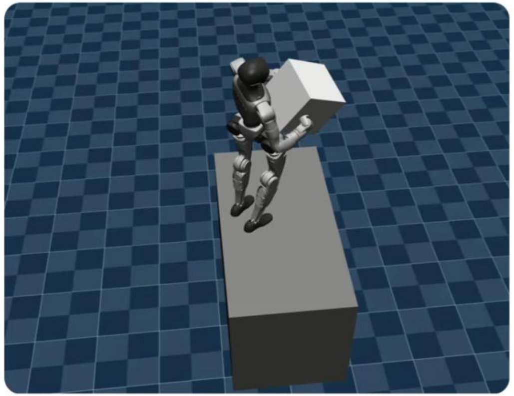
- Converts the simplified-model plan into actuator torques.
- Commonly formulated as a hierarchical program, allowing multiple tasks and constraints to be combined.
- Runs at a high control rate and compensates for tracking errors and model mismatch using feedback.



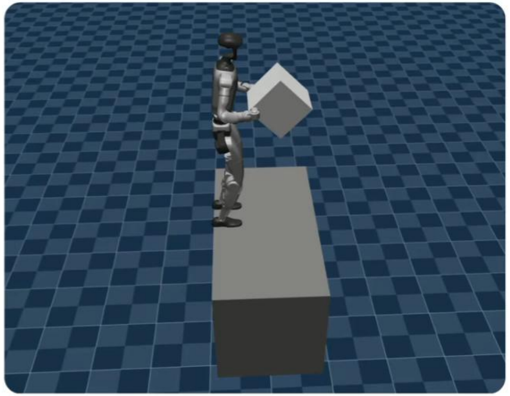
Loco-Manipulation

Contact Plan Evaluation

```
contact_mode_seq = {
    "left_sole":  [(1, 1),    (1, 1),    (1, 1),    (1, 1),    (1, 1),  (1, 1),  (),      (1, 1),  (1, 1),  (2, 1),  (2, 1)],
    "right_sole": [(1, 1),    (),        (1, 1),    (1, 1),    (1, 1),  (),      (1, 1),  (1, 1),  (),      (2, 1),  (2, 1)],
    "left_ee":    [(),        (),        (),        (9, 1),    (9, 1),  (9, 1),  (9, 1),  (9, 1),  (9, 1),  (9, 1),  (9, 1)],
    "right_ee":   [(),        (),        (),        (11, 1),   (11, 1), (11, 1), (11, 1), (11, 1), (11, 1), (11, 1), (11, 1)],
    "white_box":  [(6, 1, 1), (6, 1, 1), (6, 1, 1), (6, 1, 1), (),     (),      (),      (),      (),      (),      ()],
}
```



Sequential IK Check



Trajectory Optimization

Taouil, et al. “MotionDisco: Motion Discovery for Extreme Humanoid Loco-Manipulation,” 2026.

Learning-Based Legged Locomotion

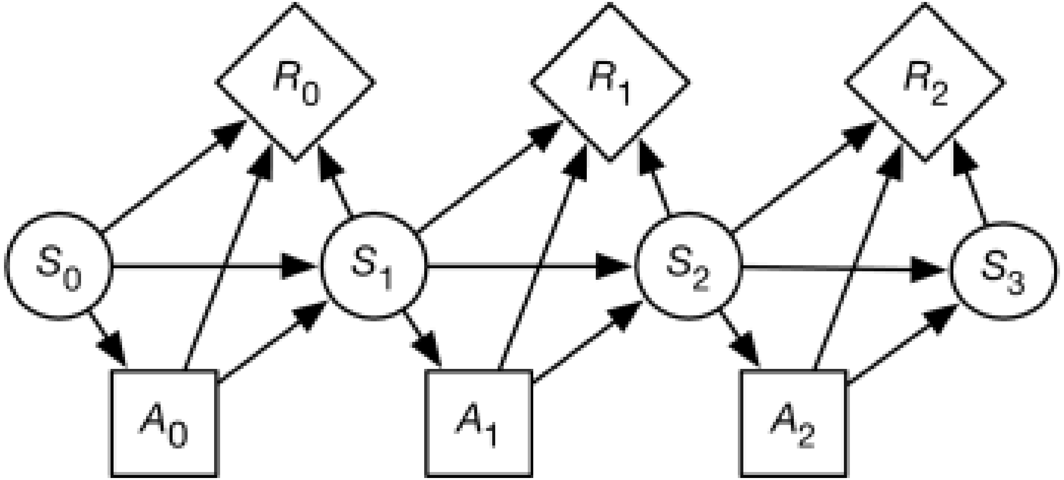
Markov Decision Process (MDP)

An **MDP** consists of the following components:

- S: A set of **states** the agent can be in
- A: A set of **actions** the agent can take
- P(s'|s, a): The **transition probability function** – the probability of reaching state s' after taking action a in state s
- R(s, a): The **reward** function – the immediate reward received after taking action a in state s
- γ (gamma): The **discount factor** – a number between 0 and 1 that determines how much future rewards are valued compared to immediate rewards

Markov Decision Process (MDP)

- The MDP satisfies the **Markov property,** which means the **next state and reward** depend only on the **current state and action**, not on the history of previous states or actions.



Reinforcement Learning Objective

- Reinforcement learning aims to maximize the expected cumulative reward over time

$$\underbrace{\mathbb{E}_{\tau \sim p_\theta(\tau)}}_{\text{Expectation over Trajectories}} \underbrace{\left[\sum_{t=1}^{T} r(s_t, a_t) \right]}_{\text{Cumulative Reward/Return}}$$

Observations and Actions

- Base linear and angular velocities
- Measurement of the gravity vector
- Joint positions and velocities
- The previous actions selected by the policy
- The 108 measurements of the terrain sampled from a grid around the robot’s base. Each measurement is the distance from the terrain surface to the robot’s base height.

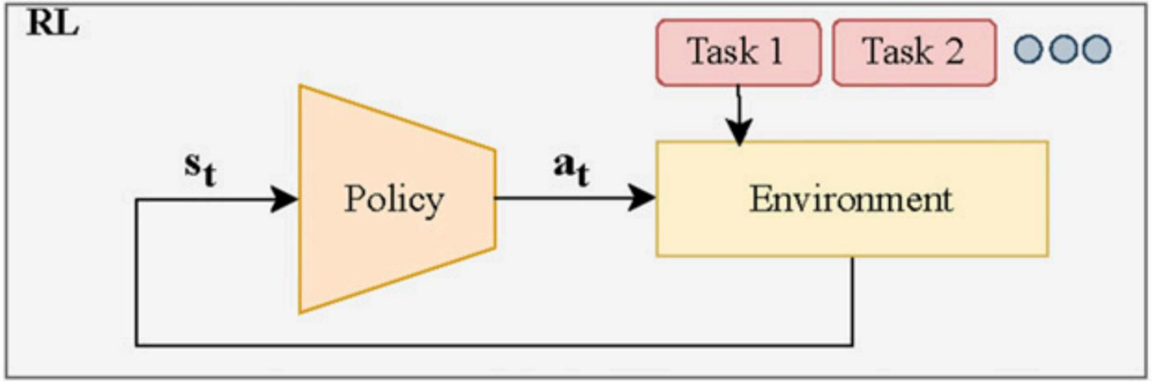
- **Action**: Desired joint positions sent to the motors. There, a PD controller produces motor torques.

Rewards

- Encourage the robot to follow the commanded velocities.
- Penalize joint torques, joint accelerations, joint target changes, and collisions, to create a smoother, more natural motion.
- Contacts with the knees, shanks or between the feet and a vertical surface are considered collisions.
- Contacts with the base are considered crashes and lead to resets.
- Reward term encouraging the robot to take longer steps, which results in a more visually appealing behavior.

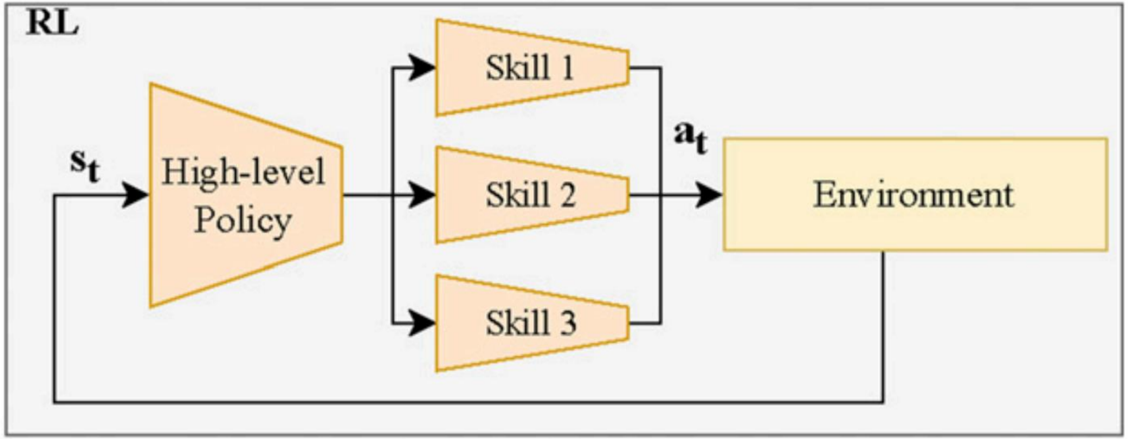
Learning Frameworks

- **Curriculum learning**: it creates progressively challenging learning milestones to enhance the policy



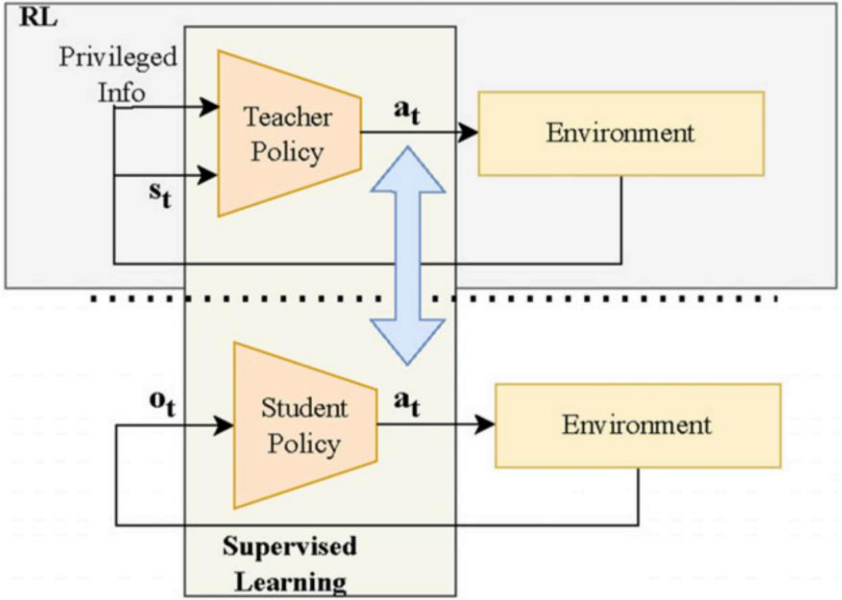
Learning Frameworks

- **Hierarchical learning**: it is suitable for long horizon problems like navigation, it decomposes the problem into hierarchies



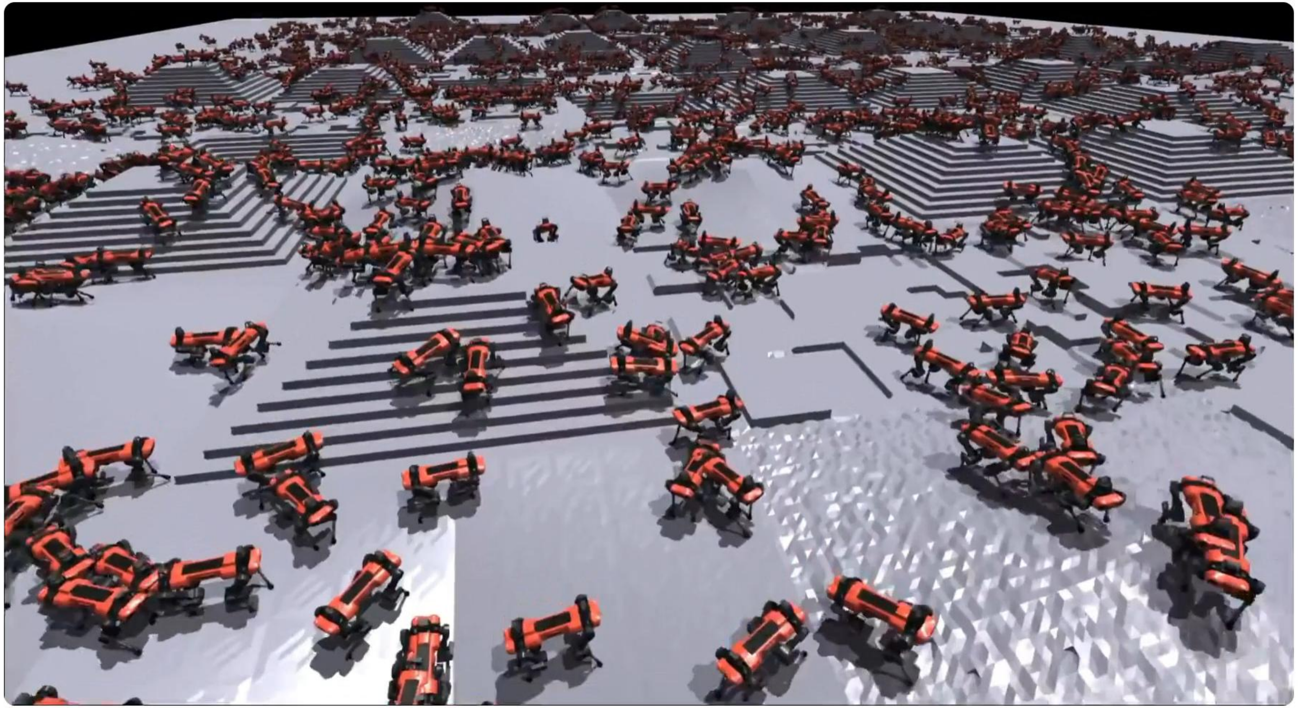
Learning Frameworks

- **Privileged learning**: it addresses the partial observability problem in robotics.



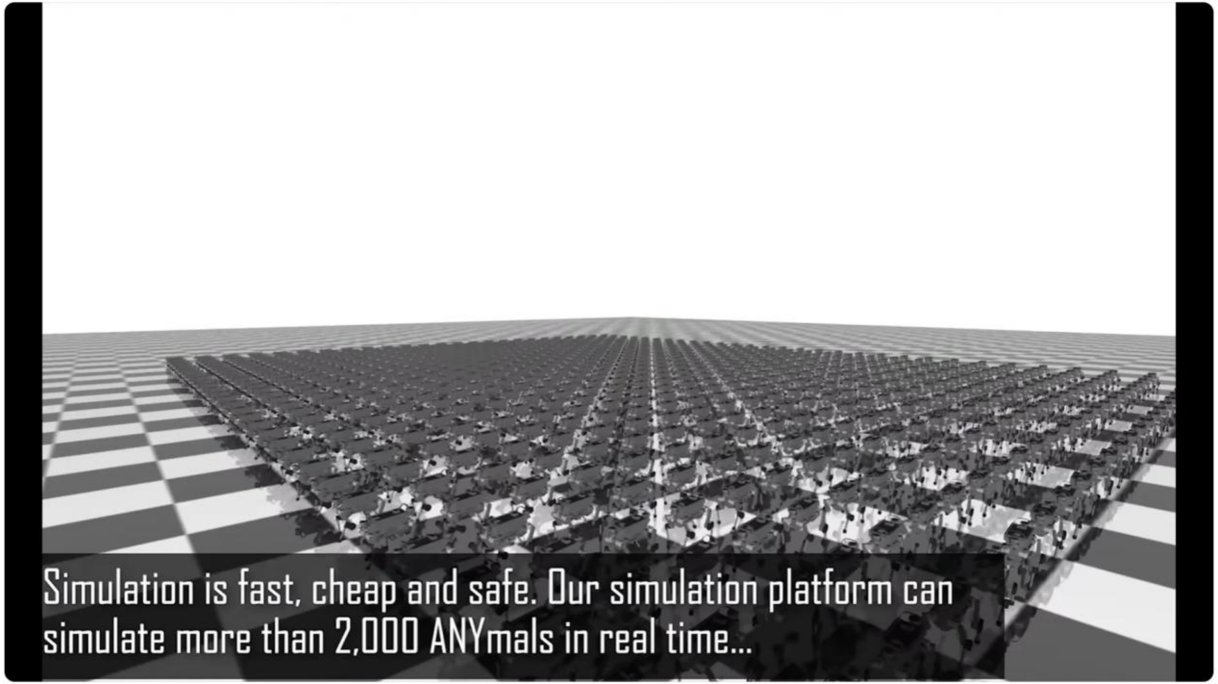
Sehoon Ha, et al. “Learning-based legged locomotion: State of the art and future perspectives,” ijrr, 2025.

Learning Locomotion of Quadrupeds



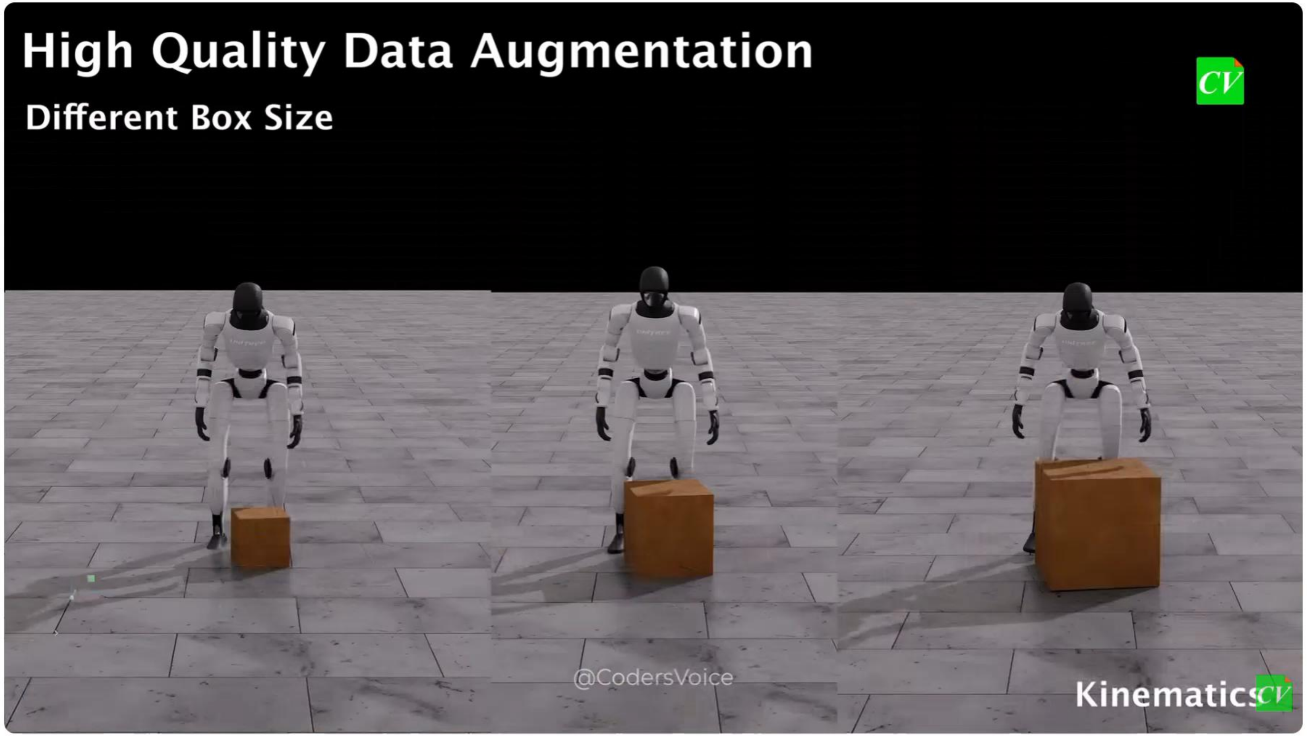
Rudin, Nikita, et al. “Learning to walk in minutes using massively parallel deep reinforcement learning,” PMLR, 2022.

Sim to Real Transfer



Hwangbo, et al. “Learning Agile and Dynamic Motor Skills for Legged Robots,” Science Robotics, 2019.

Loco-Manipulation



Yang, et al. “OmniRetarget: Interaction-Preserving Data Generation for Humanoid Whole-Body Control,” 2026.

Literature

- Wensing , et al. “Optimization-based control for dynamic legged robots,” IEEE TRO, 2024.
- Sehoon Ha, et al. "Learning-based legged locomotion: State of the art and future perspectives," ijrr, 2025
- Sutton, Richard S., and Andrew G. Barto. “Reinforcement learning: An introduction,” Vol. 1. No. 1. Cambridge: MIT press, 1998.
- Rudin, Nikita, et al. "Learning to walk in minutes using massively parallel deep reinforcement learning." Conference on Robot Learning. PMLR, 2022.
- Victor Dh´Edin, et al. “Diffusion-based learning of contact plans for agile locomotion,” Humanoids, 2024.
- Sleiman, et al. “A Unified MPC Framework for Whole-Body Dynamic Locomotion and Manipulation,” IEEE ICRA, 2021.
- Taouil, et al. “MotionDisco: Motion Discovery for Extreme Humanoid Loco-Manipulation,” 2026.
- Hwangbo, et al. “Learning Agile and Dynamic Motor Skills for Legged Robots,” Science Robotics, 2019.
- Yang, et al. “OmniRetarget: Interaction-Preserving Data Generation for Humanoid Whole-Body Control,” 2026.