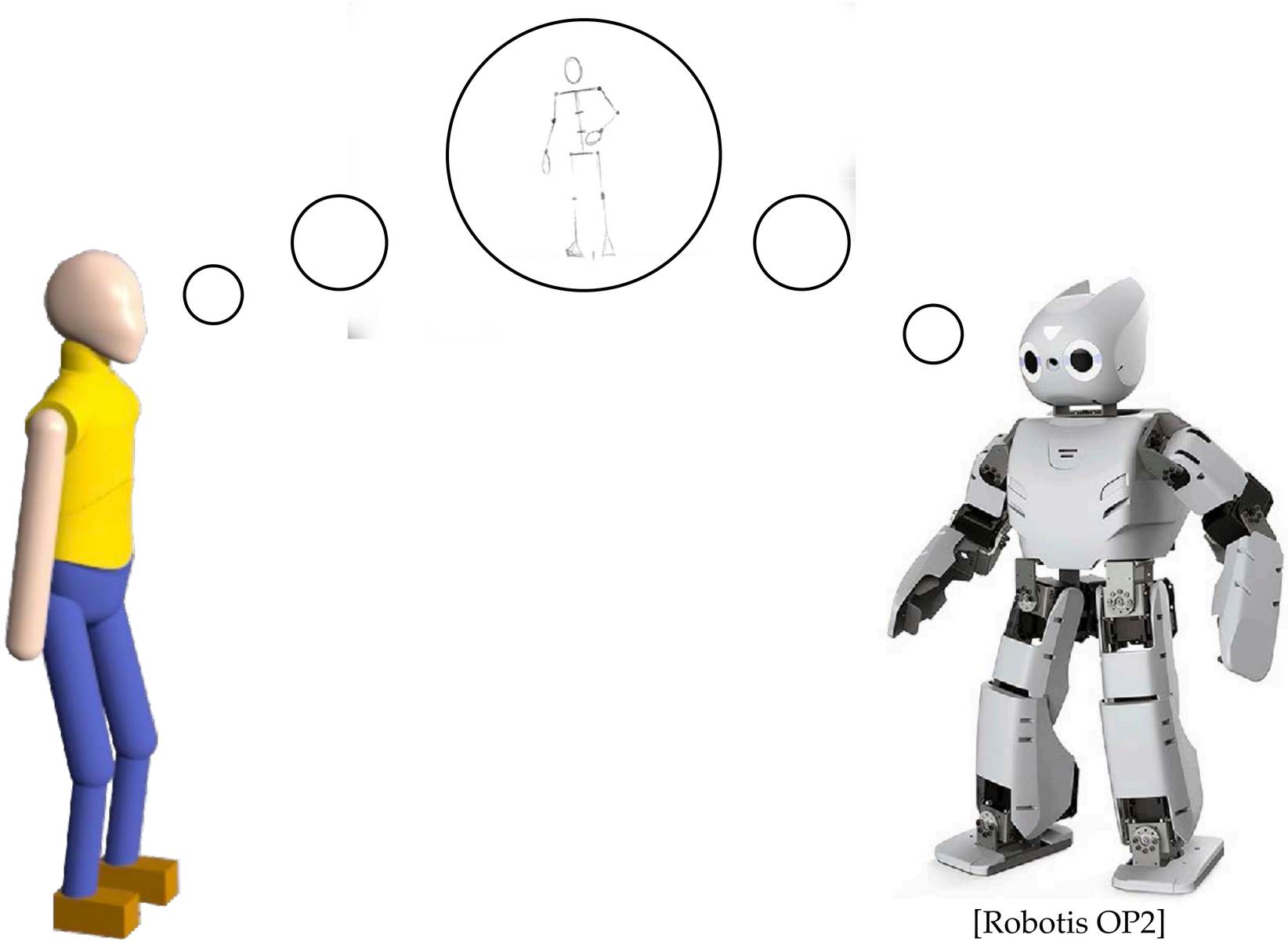
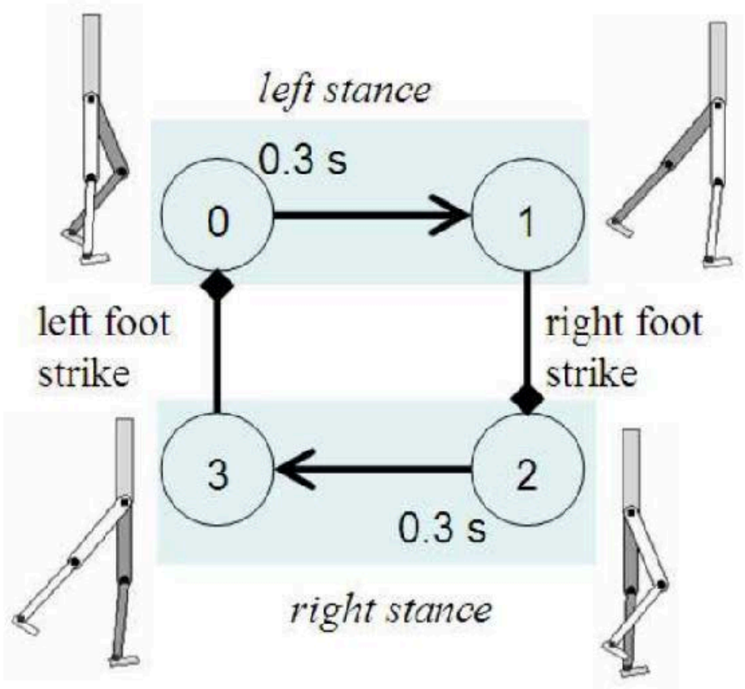
Learning Symmetric and Low-energy Locomotion

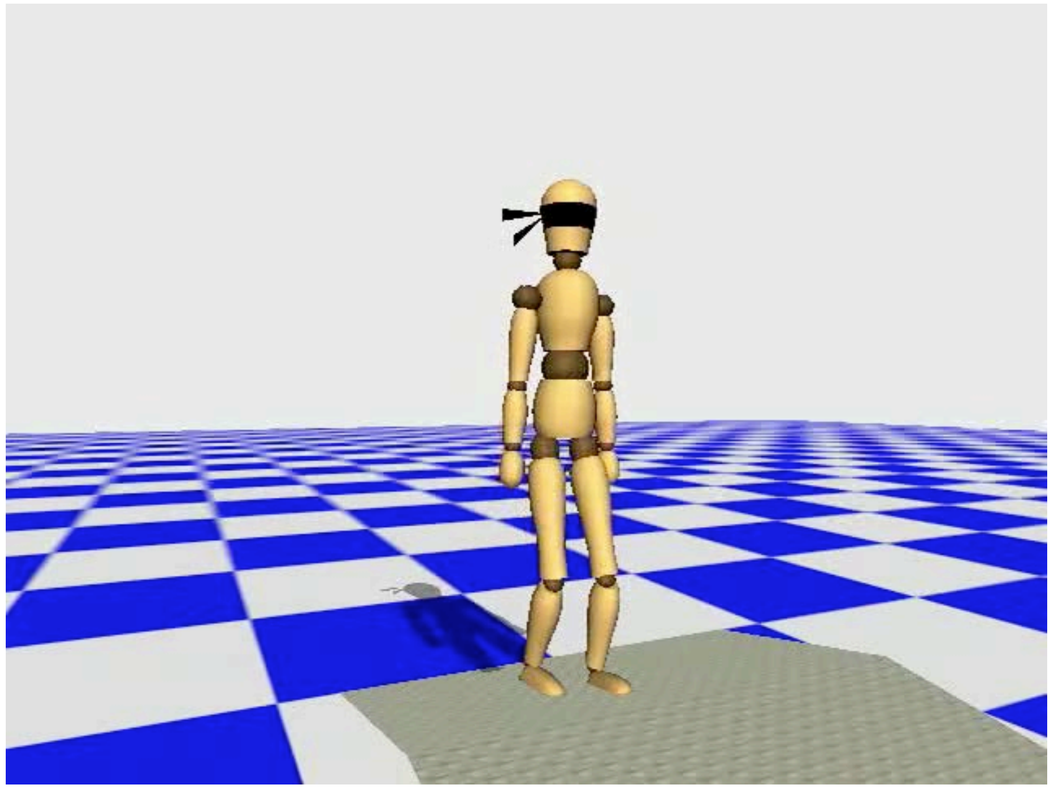
Wenhao Yu, Greg Turk, C. Karen Liu
Georgia Institute of Technology

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3591461/
https://www.youtube.com/watch?v=Wz3beWec7D4



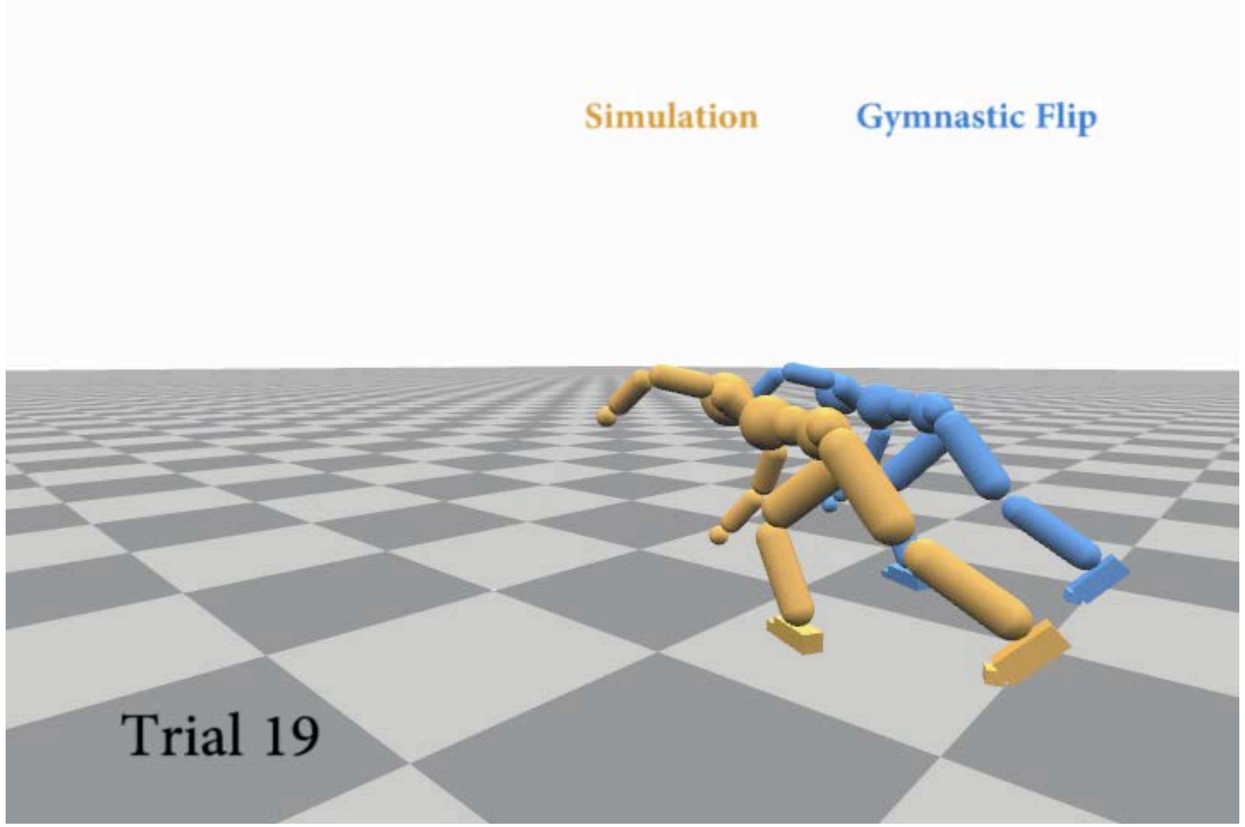
[Robotis OP2]





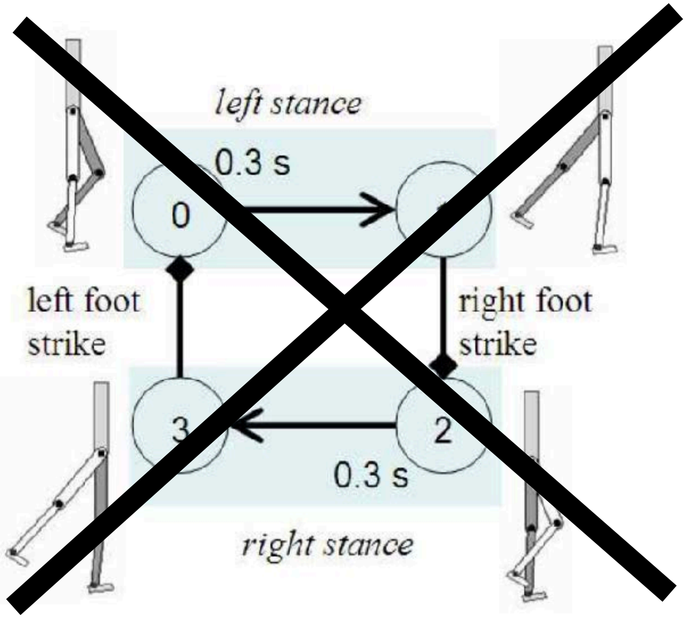
[Yin et al. 2009]

[OptiTrack]

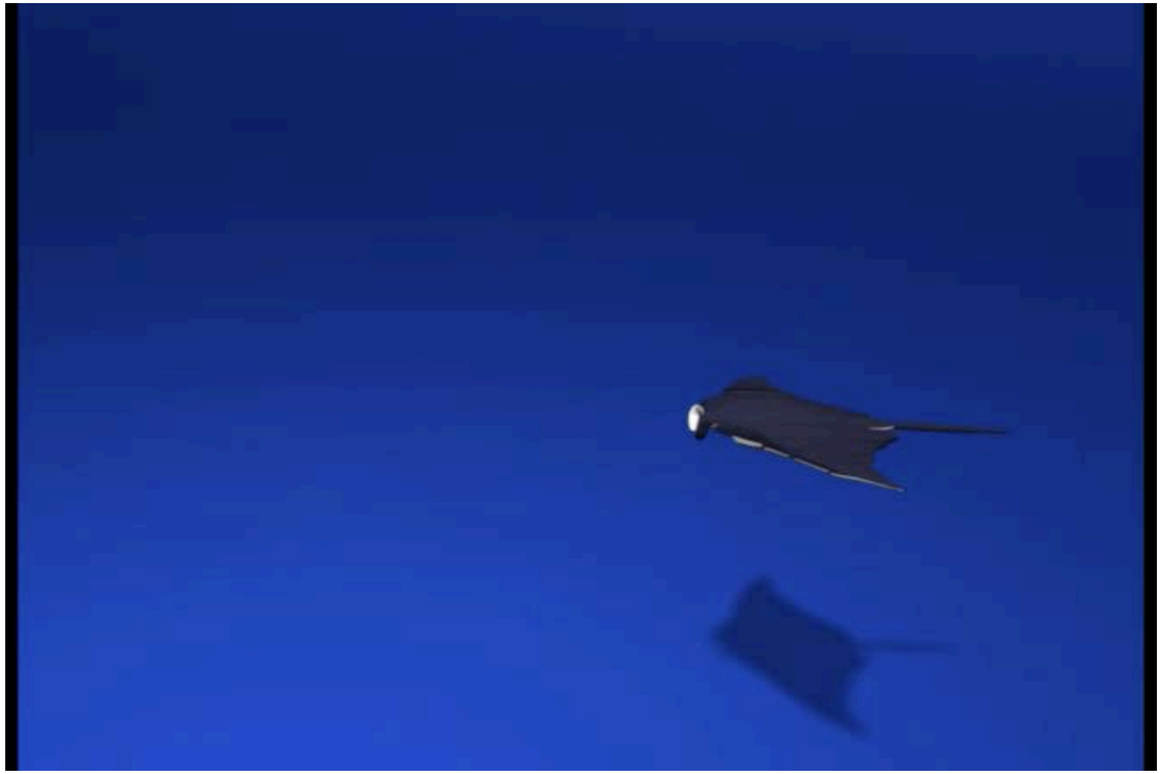


[Liu et al. 2015]

No FSM

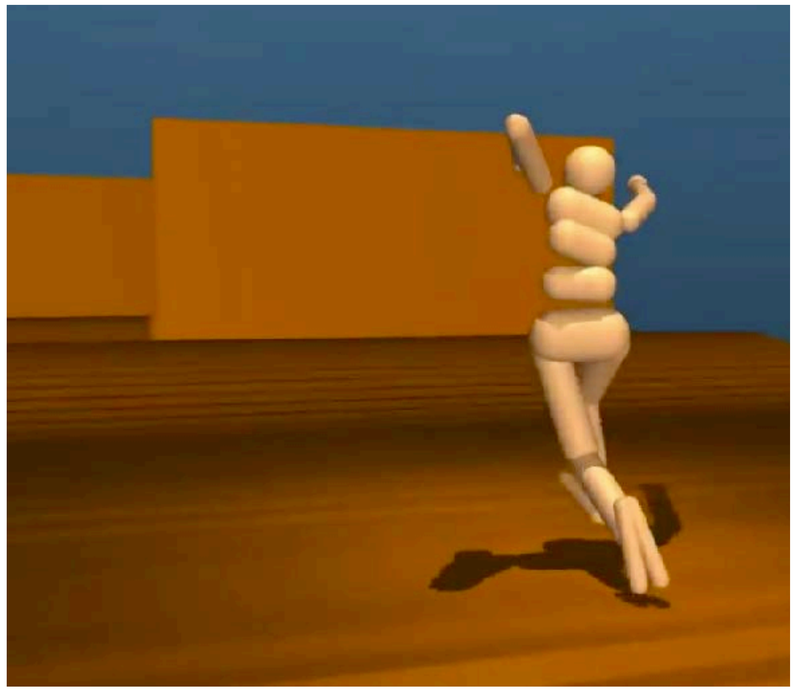


No Mocap

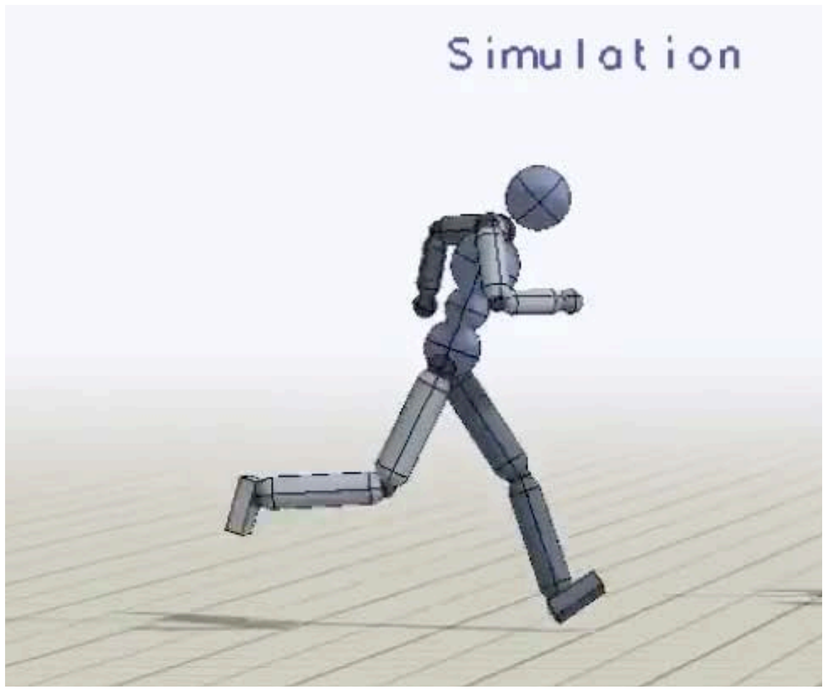


[Tan et al. 2011]

[Tan et al. 2014]

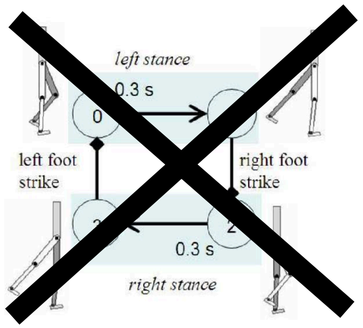


[Heess et al. 2017]

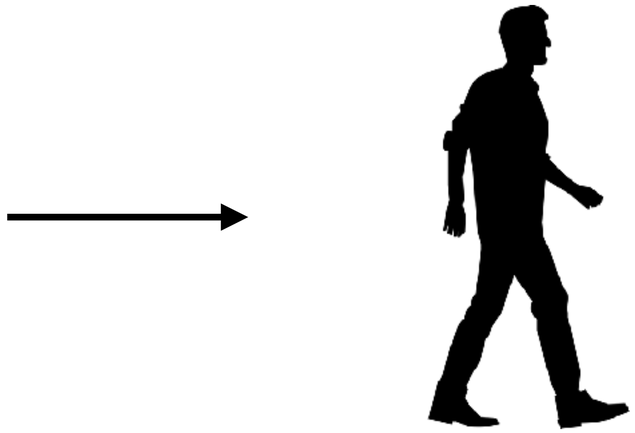


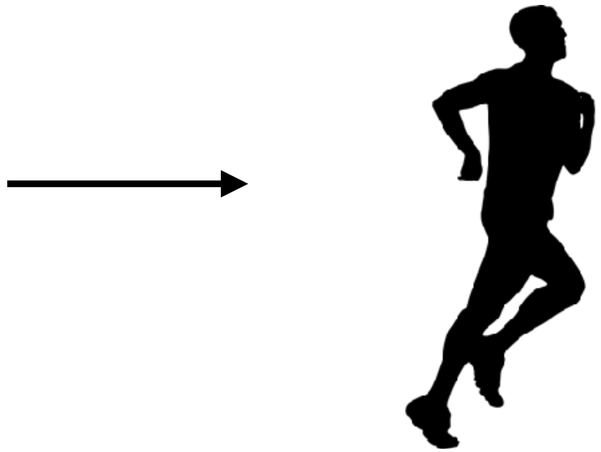
[Peng et al. 2018]

Deep Reinforcement Learning

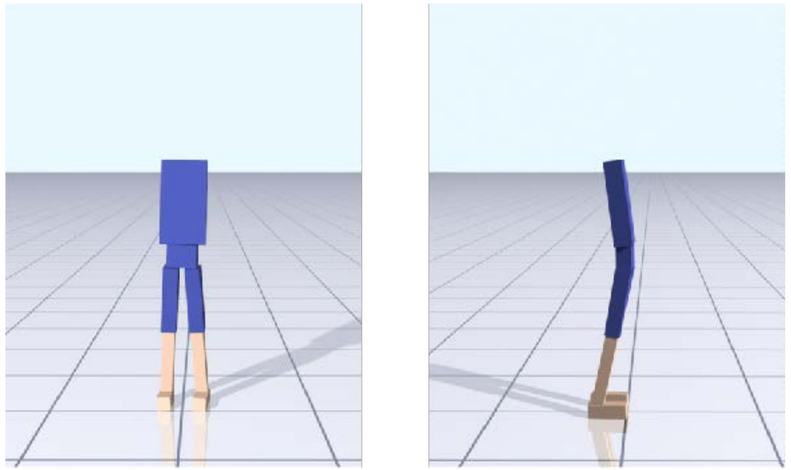


Deep Reinforcement Learning

1.0 m/s 

3.0 m/s 

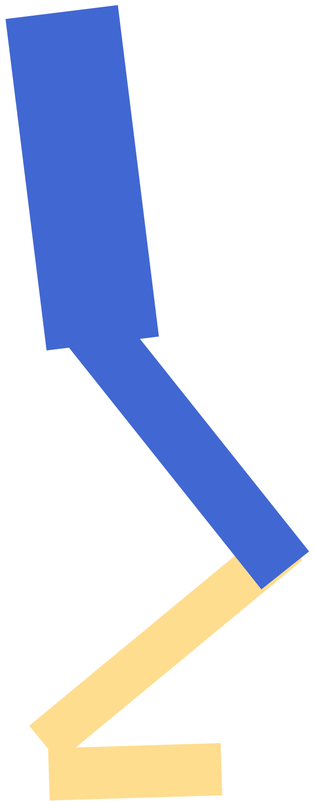

Deep Reinforcement Learning
Energy Minimization
Gait Symmetry

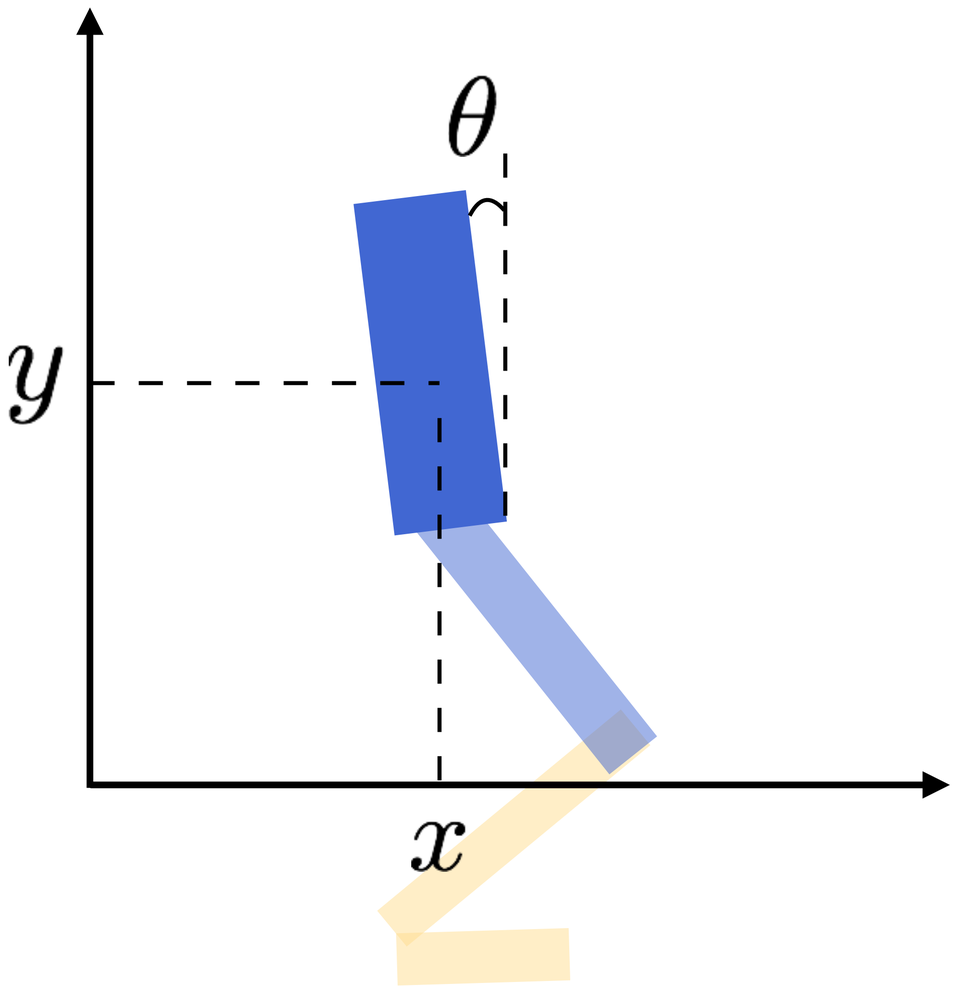


Deep Reinforcement Learning

?

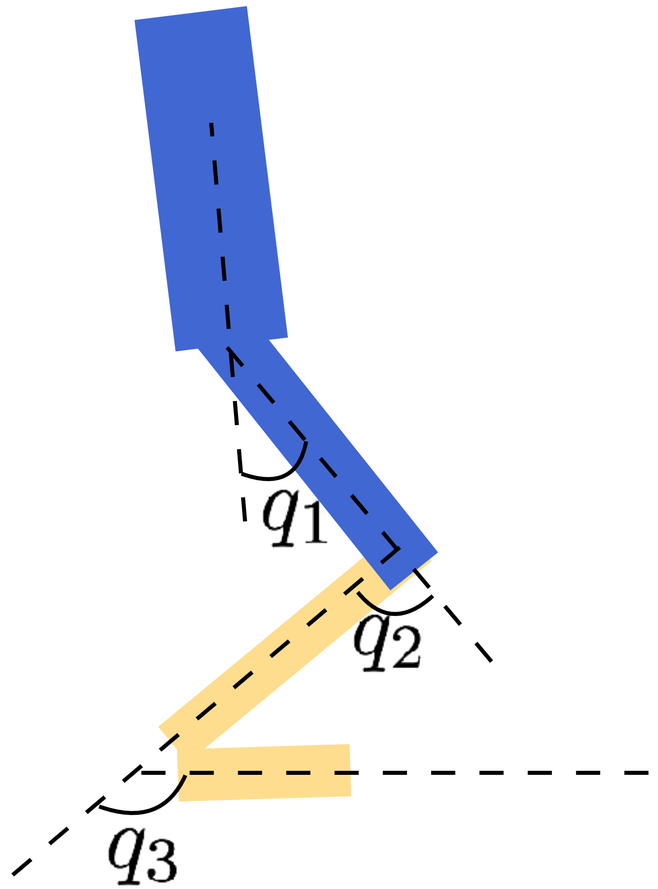
State:



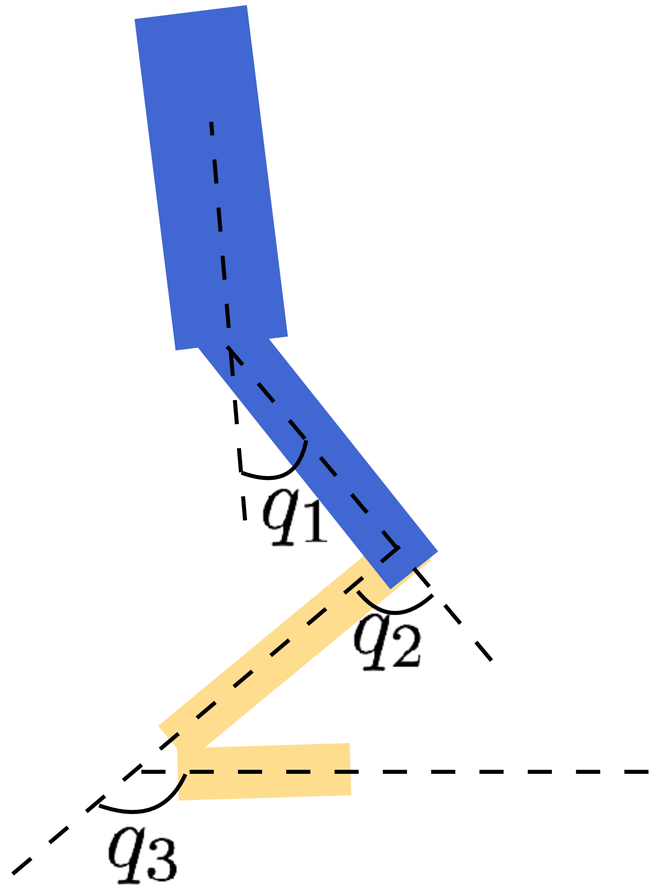


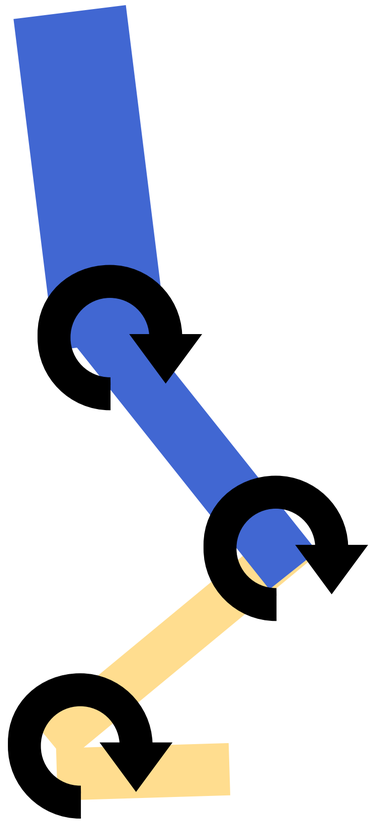
State: $\mathbf{s} = [x, y, \theta$

State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3]$



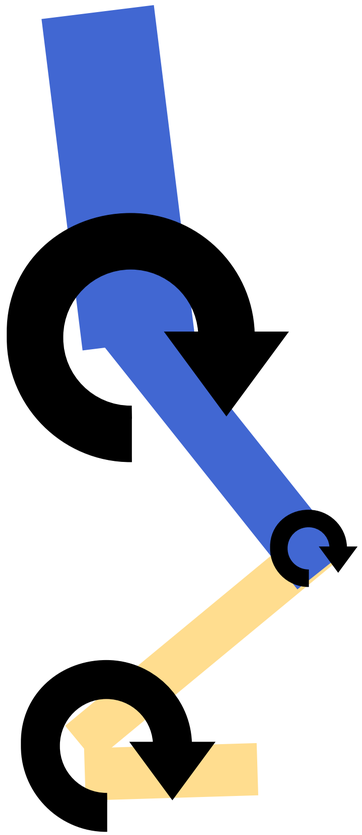
State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3, \dot{x}, \dot{y}, \dot{\theta}, \dot{q_1}, \dot{q_2}, \dot{q_3}]$





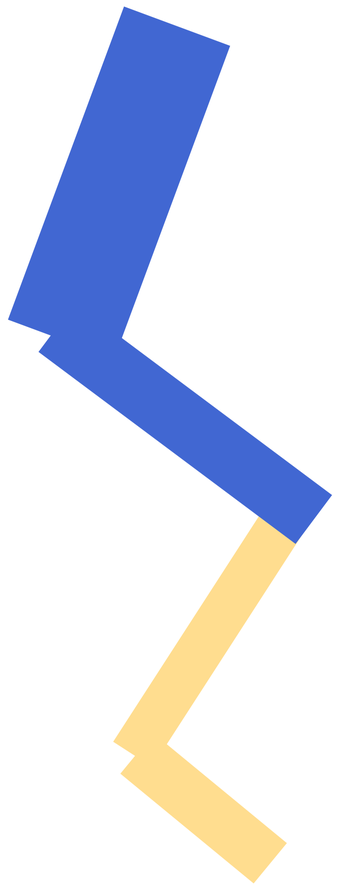
State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3, \dot{x}, \dot{y}, \dot{\theta}, \dot{q_1}, \dot{q_2}, \dot{q_3}]$

Action: $\mathbf{a} = [a_1, a_2, a_3]$



State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3, \dot{x}, \dot{y}, \dot{\theta}, \dot{q_1}, \dot{q_2}, \dot{q_3}]$

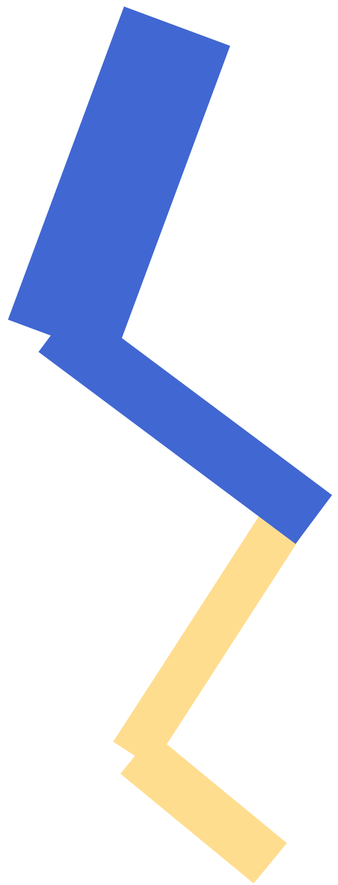
Action: $\mathbf{a} = [a_1, a_2, a_3]$



State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3, \dot{x}, \dot{y}, \dot{\theta}, \dot{q_1}, \dot{q_2}, \dot{q_3}]$

Action: $\mathbf{a} = [a_1, a_2, a_3]$

Transition: $\mathbf{s^{t+1}} = \mathbf{T}(\mathbf{s^t}, \mathbf{a})$



State: $\mathbf{s} = [x, y, \theta, q_1, q_2, q_3, \dot{x}, \dot{y}, \dot{\theta}, \dot{q_1}, \dot{q_2}, \dot{q_3}]$

Action: $\mathbf{a} = [a_1, a_2, a_3]$

Transition: $\mathbf{s^{t+1}} = \mathbf{T}(\mathbf{s^t}, \mathbf{a})$

Reward: $r = R(\mathbf{s}, \mathbf{a}) = x - \sum_{i=1}^{3} |a_i|$

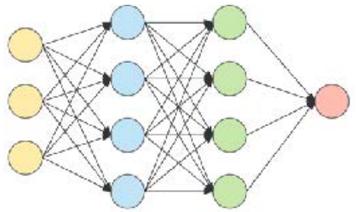
Markov Decision Process: $\{\ \mathbf{s} \quad \mathbf{a} \quad \mathbf{T} \quad R\ \}$

Control Policy: $\mathbf{a}_t \sim \pi(\mathbf{a}|\mathbf{s}_t)$

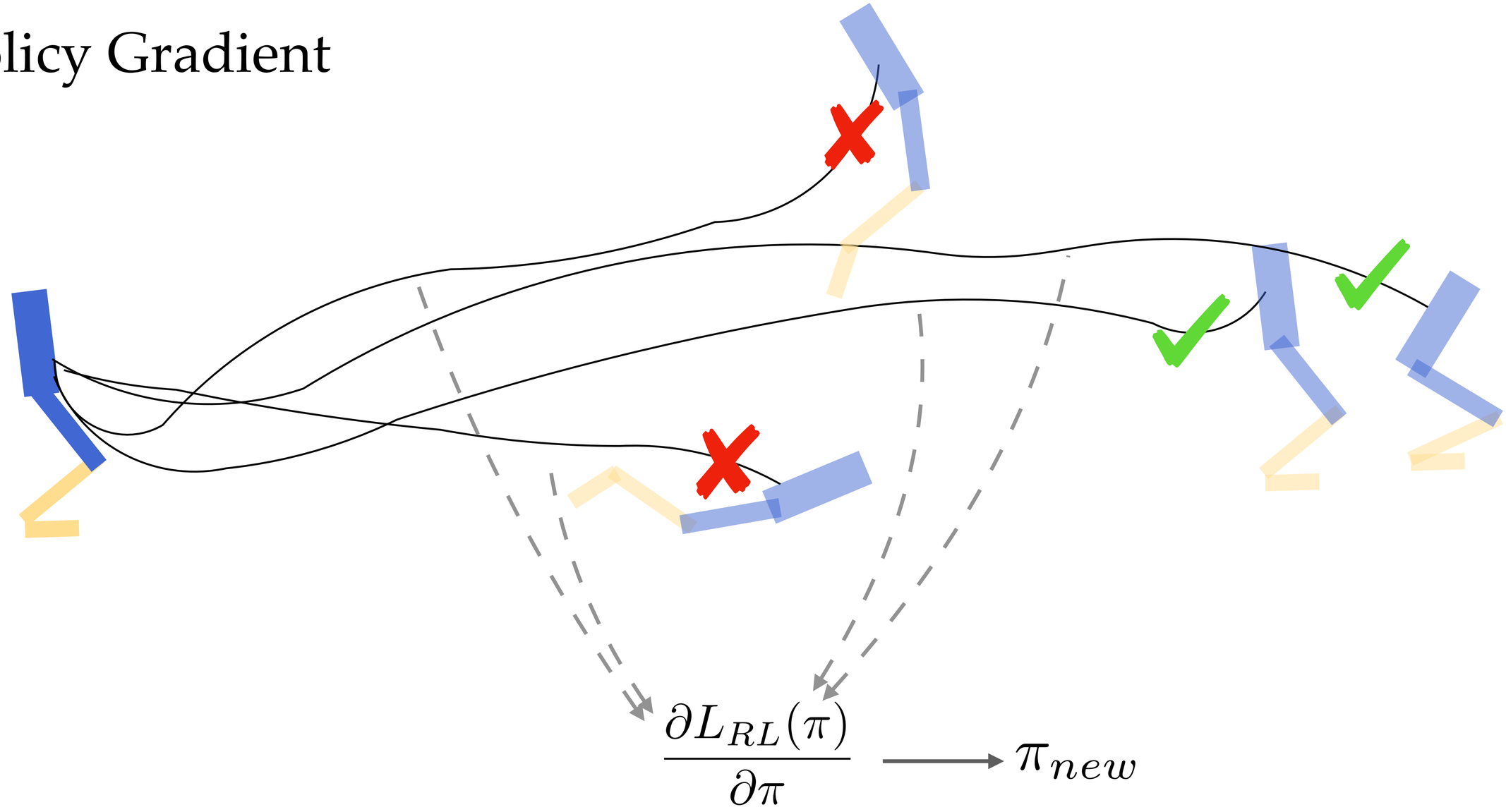
Rollout: $[\mathbf{s}_0, \mathbf{a}_0, \mathbf{s}_1, \mathbf{a}_1, \ldots, \mathbf{s}_T]$

Reinforcement Learning: $\text{maximize } L_{RL}(\pi) = \mathbb{E}[\sum_{t=1}^{T} R(\mathbf{s}^{t+1}, \mathbf{a}^t)]$

DeepReinforcement Learning: π :



Policy Gradient



$$\mathbf{s} = [\mathbf{q}, \dot{\mathbf{q}}, \bar{v}, contact]$$

$$\mathbf{a} = [a_{0-14}]$$

$$\mathbf{T}(\cdot) =$$ DART Dynamic Animation and Robotics Toolkit

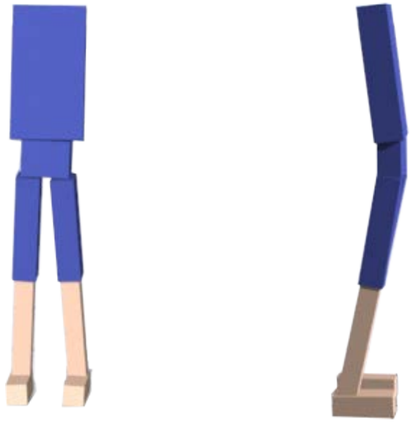
$$R(\mathbf{s}, \mathbf{a}) = -||v - \bar{v}|| \quad *w_v$$

$$-||\mathbf{a}||_1 \quad *w_a \leftarrow$$

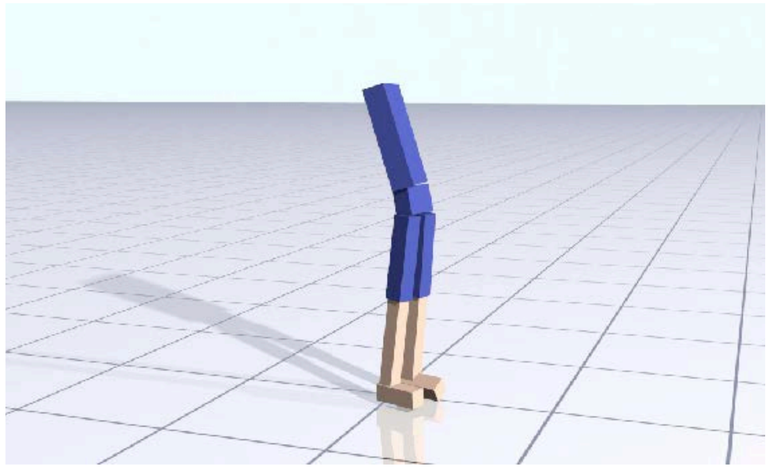
$$-\text{LaternalDeviation} \quad *w_L$$

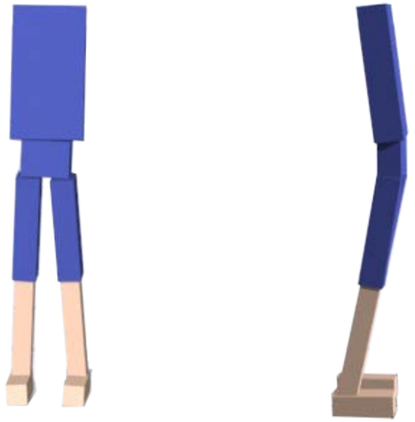
$$-\text{TorsoRotation} \quad *w_T$$

$$+\text{AliveBonus}$$



Deep Reinforcement Learning





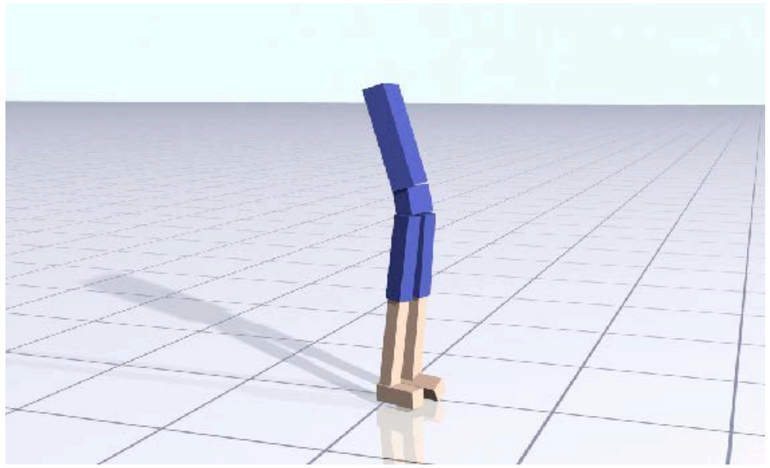
Curriculum Learning

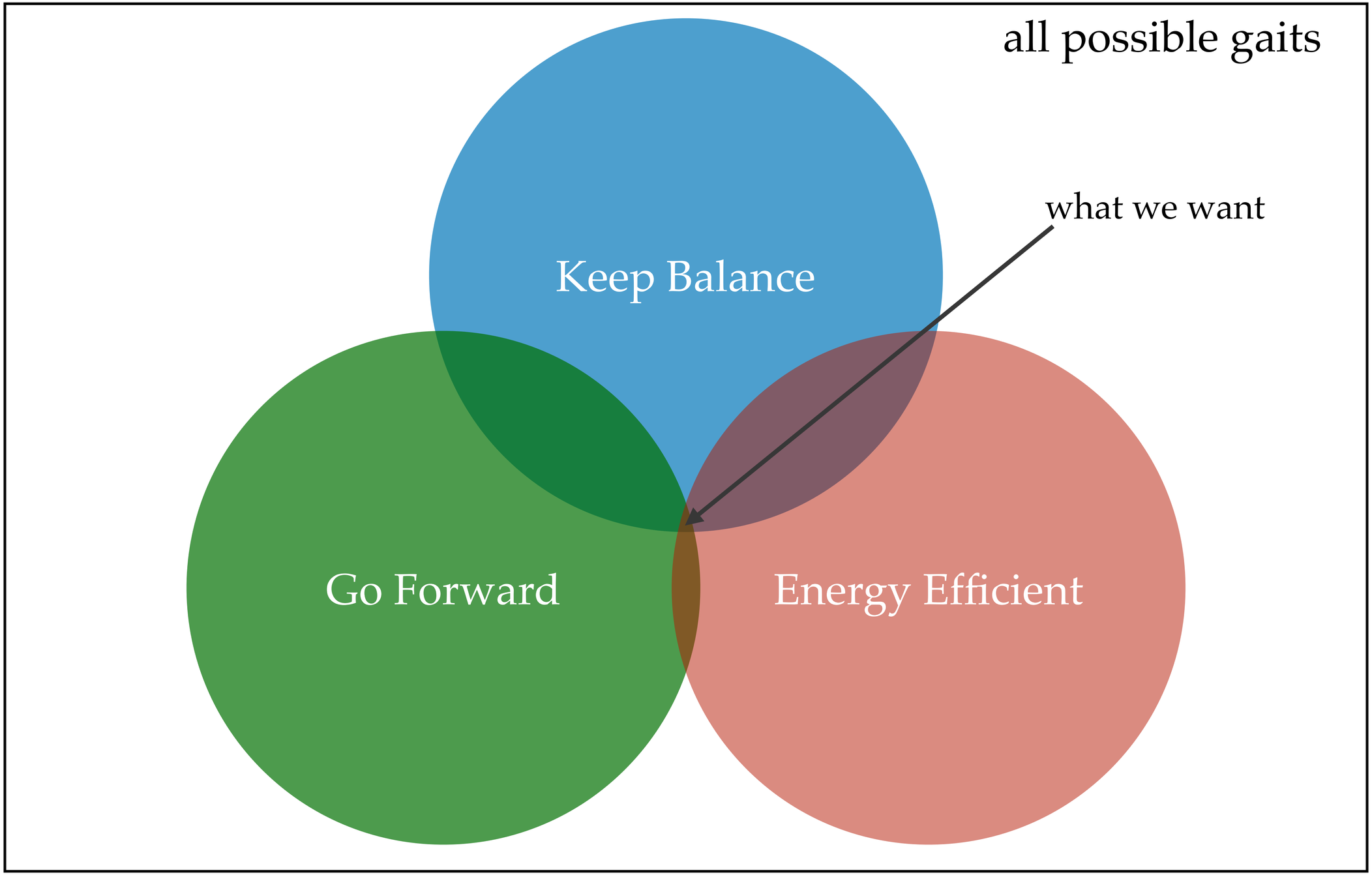
Deep Reinforcement Learning

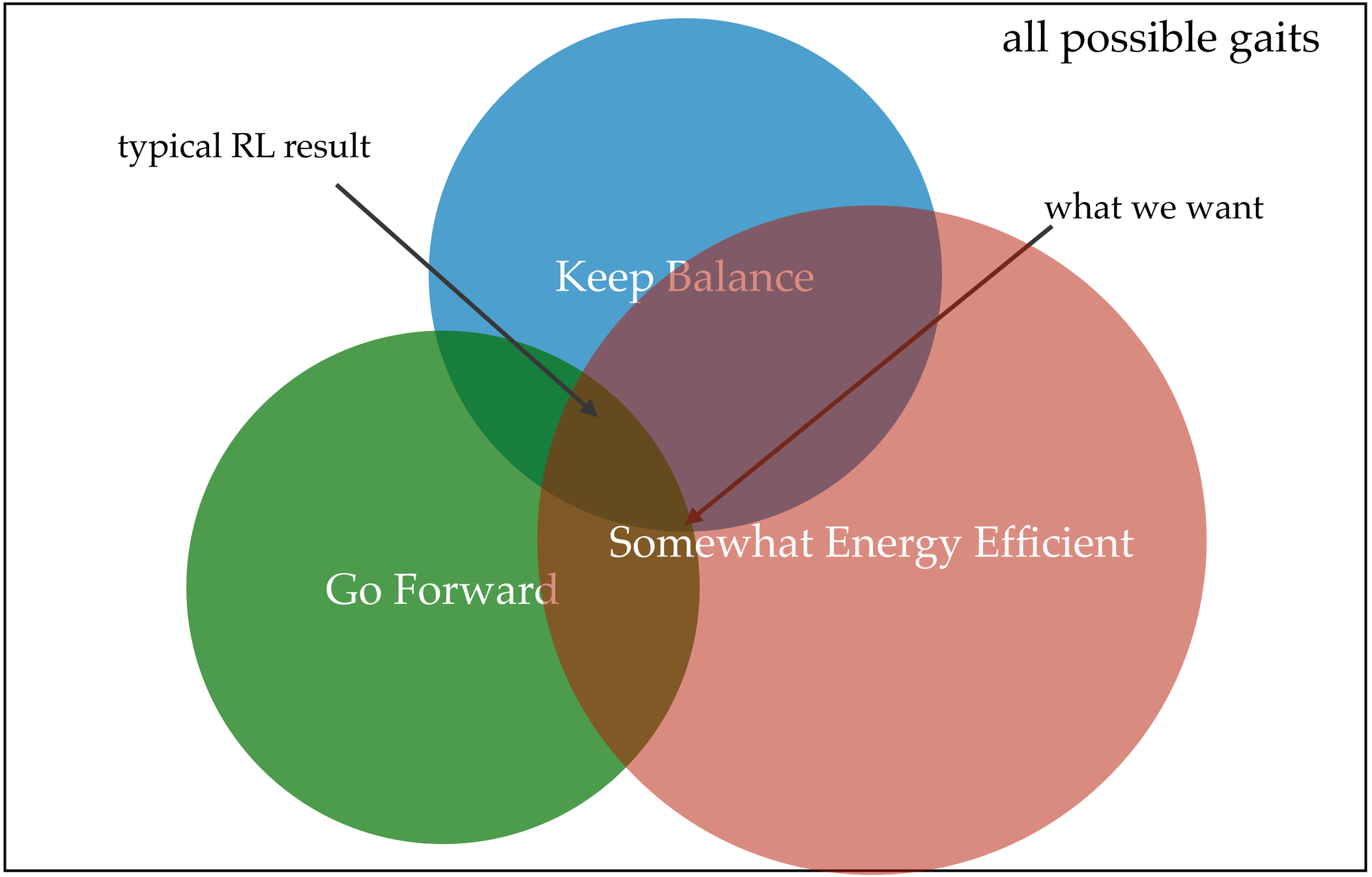
+

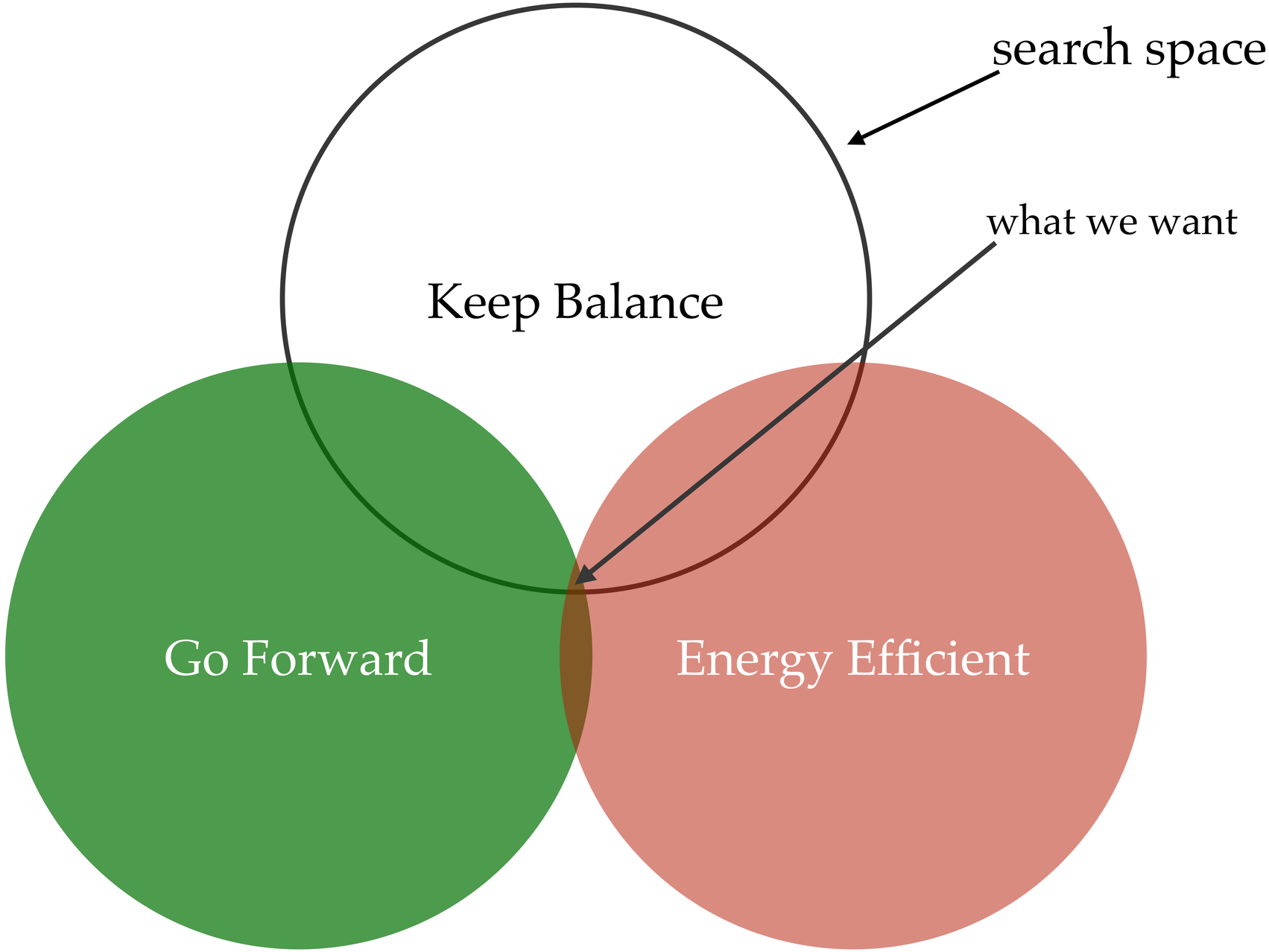
Energy Minimization

?




all possible gaits
what we want
Keep Balance
Go Forward
Energy Efficient


all possible gaits
typical RL result
what we want
Keep Balance
Somewhat Energy Efficient
Go Forward

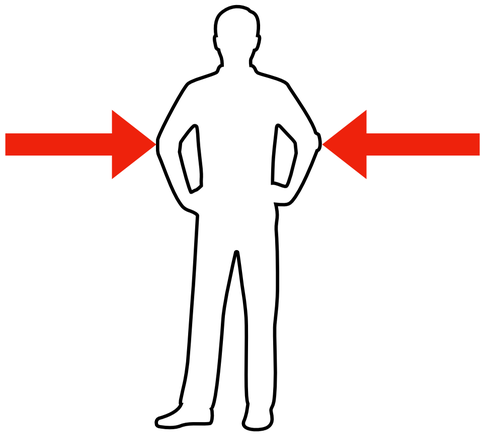

search space
what we want
Keep Balance
Go Forward
Energy Efficient

Keep Balance

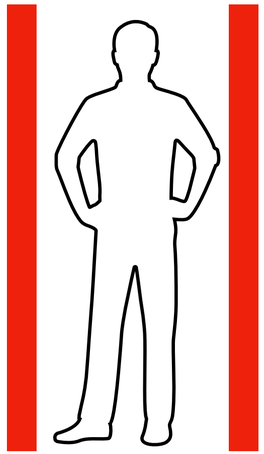
Go Forward

Energy Efficient

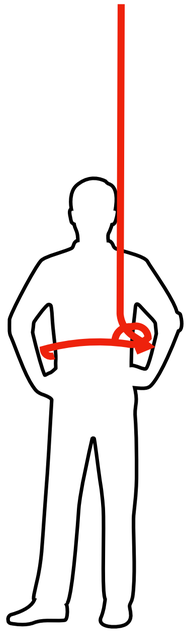
https://www.youtube.com/watch?v=5BiesFPqYWE
https://www.youtube.com/watch?v=5a21xnaqCzk
https://www.youtube.com/watch?v=LjnZHbFpDxc



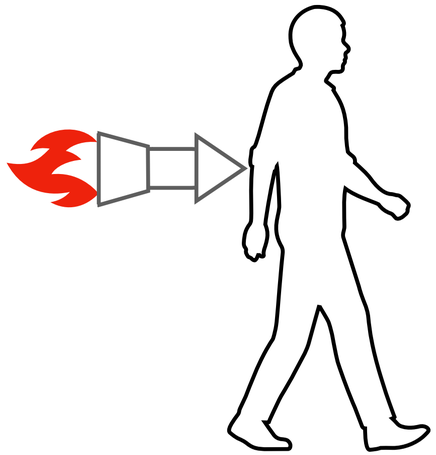
lateral push



walls

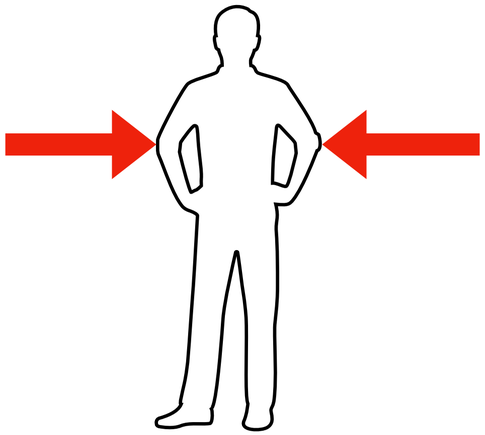


waist support



propel forward

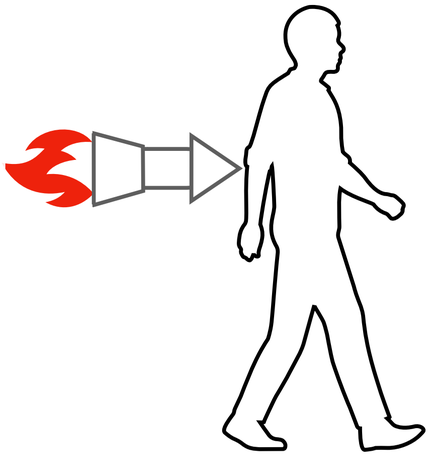
treadmill



lateral push

walls

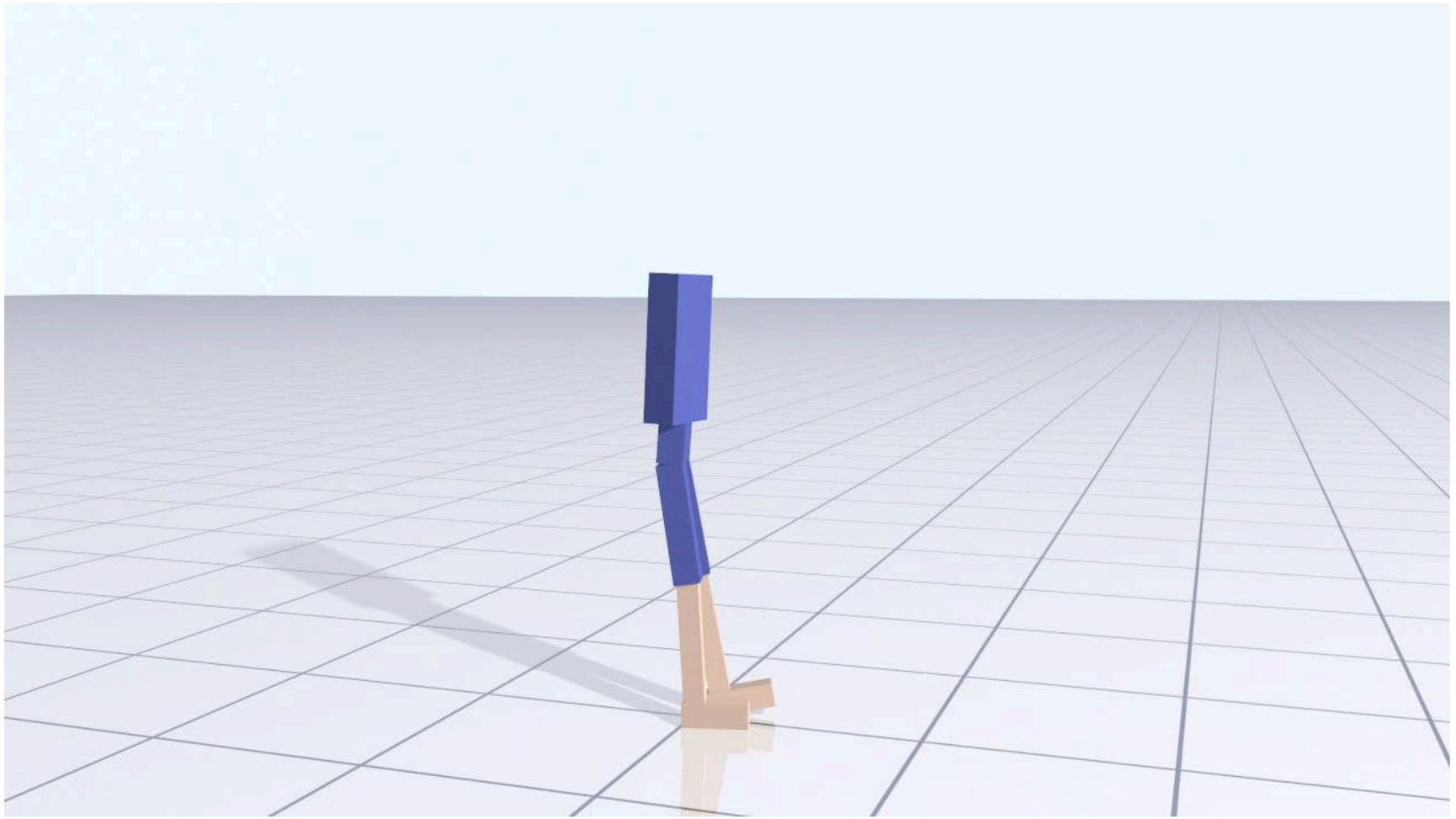
waist support



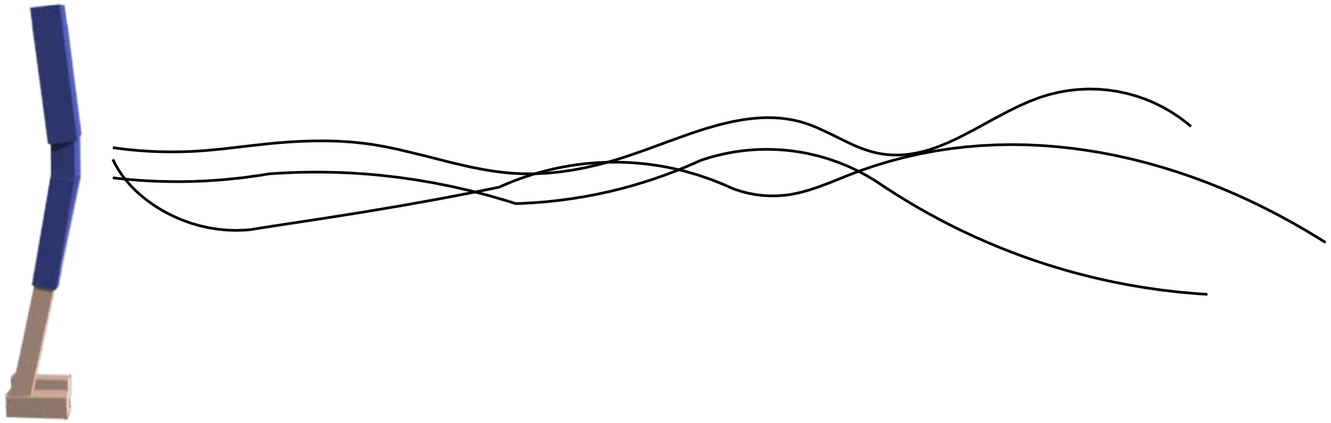
propel forward

treadmill

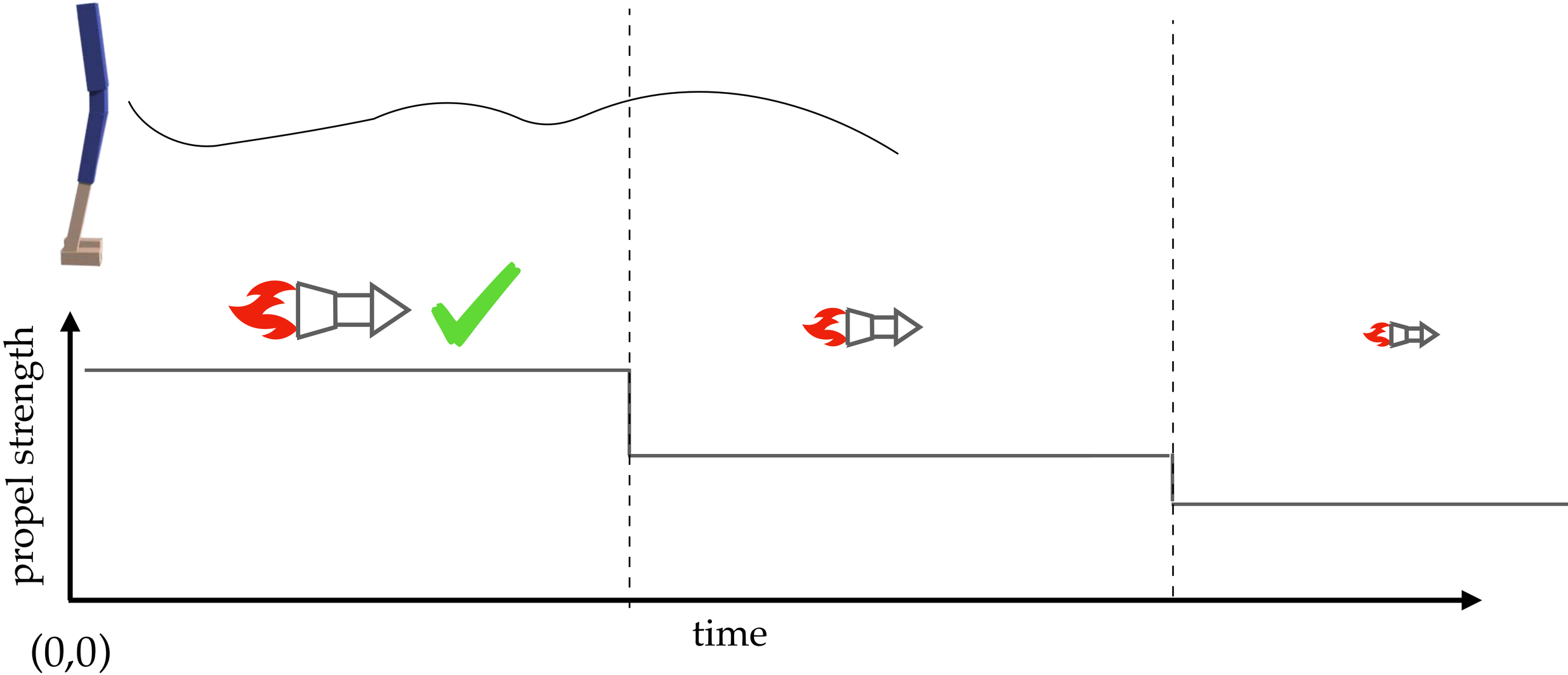
With Assistance



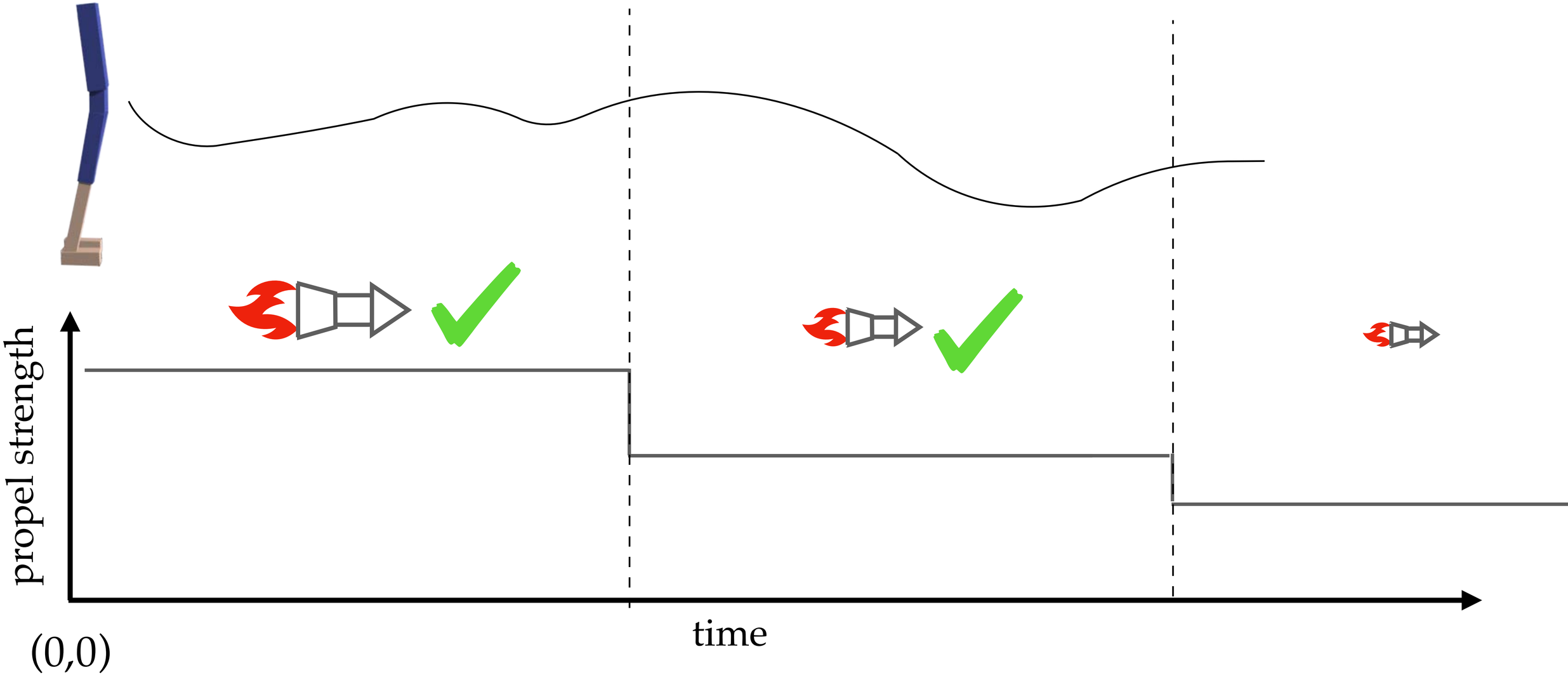
Iteration i



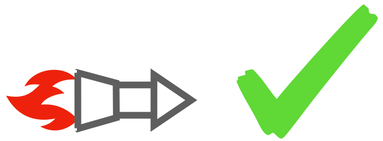
Iteration i

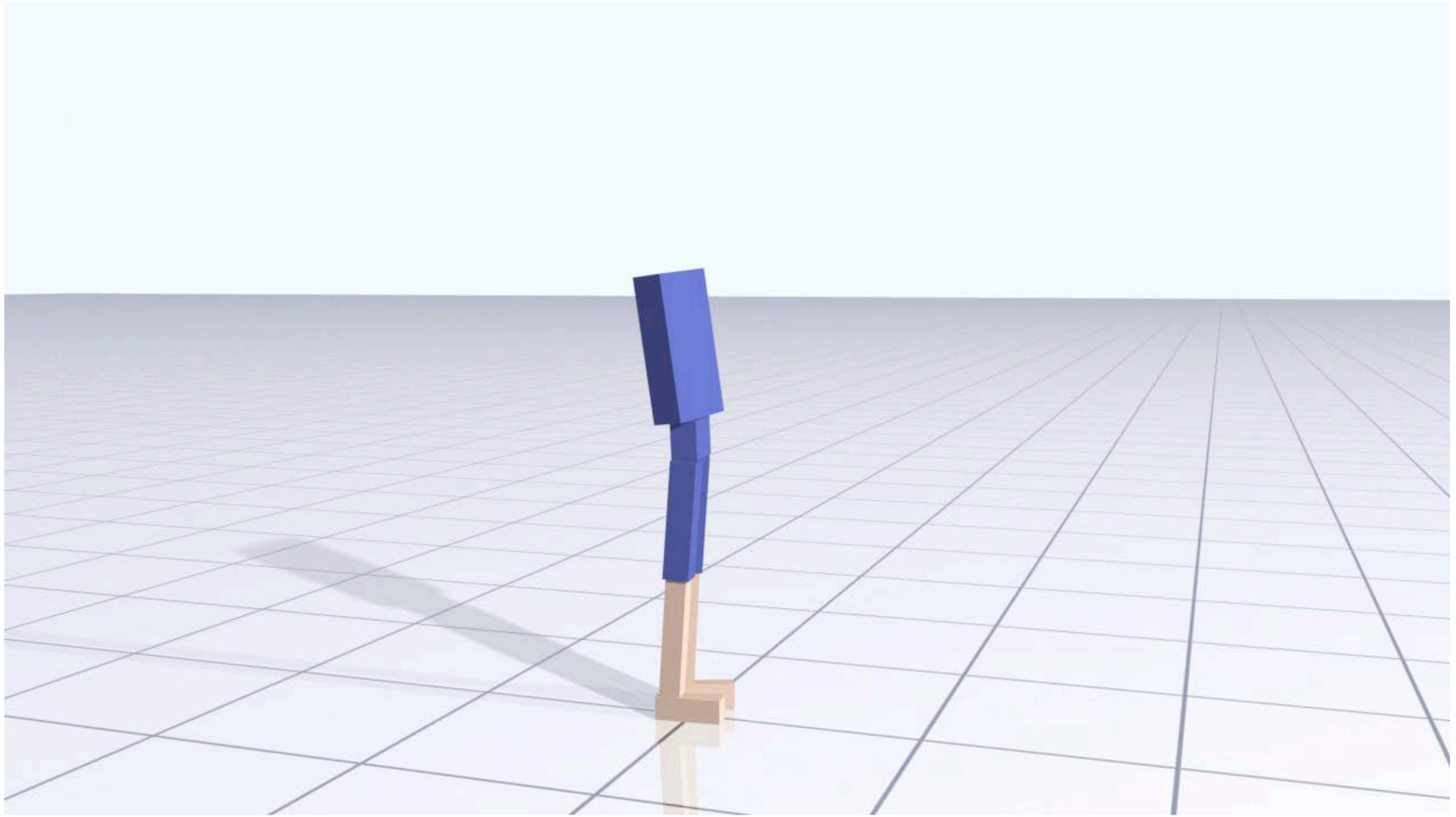


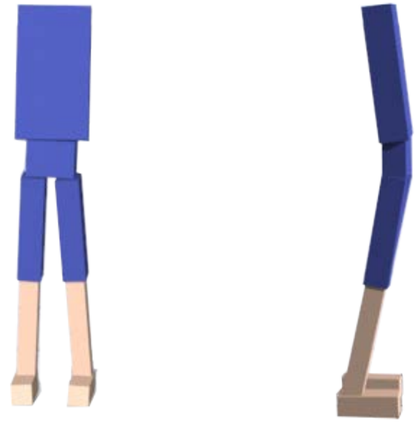
Iteration i



Iteration i+1







Curriculum Learning

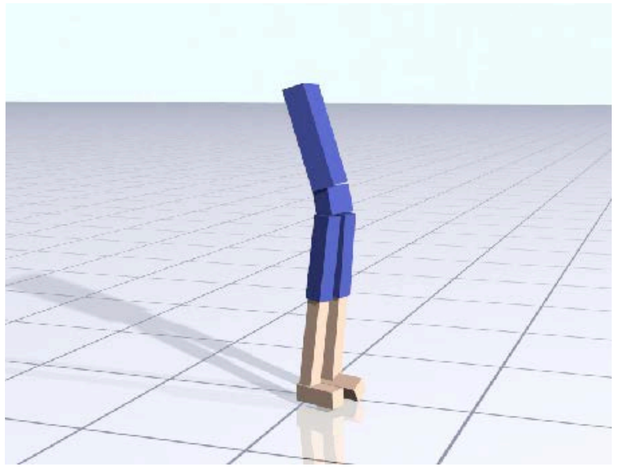
Deep Reinforcement Learning

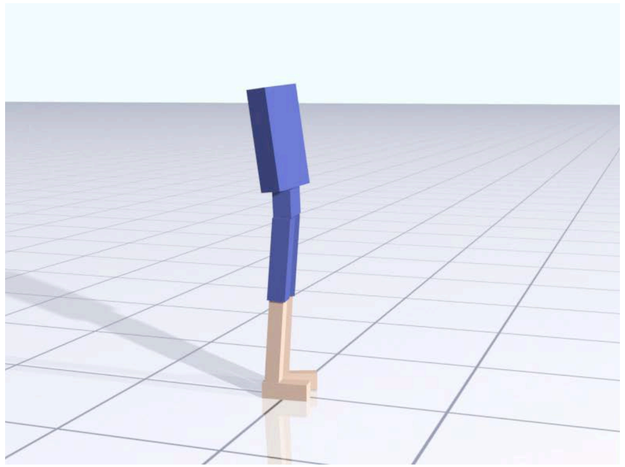
+

Energy Minimization

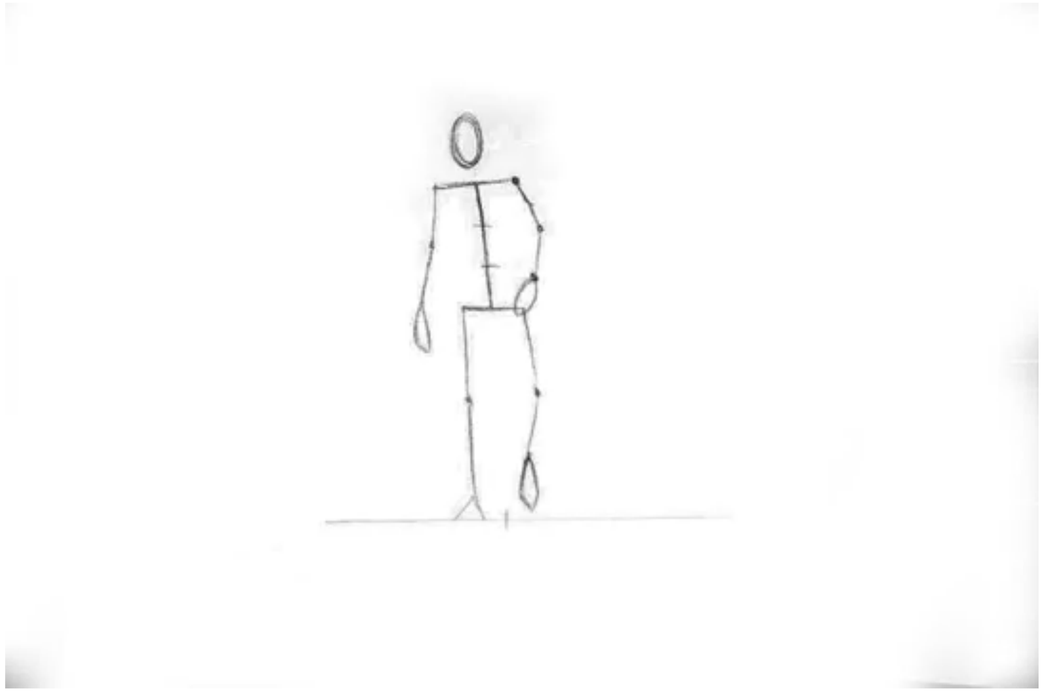
+

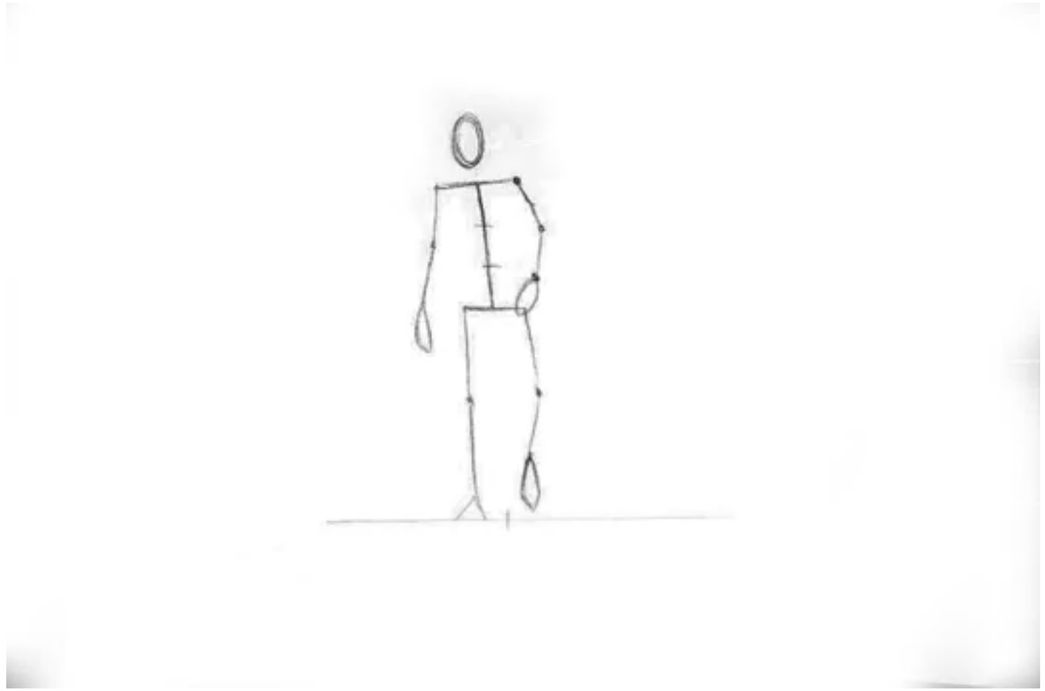
Gait Symmetry



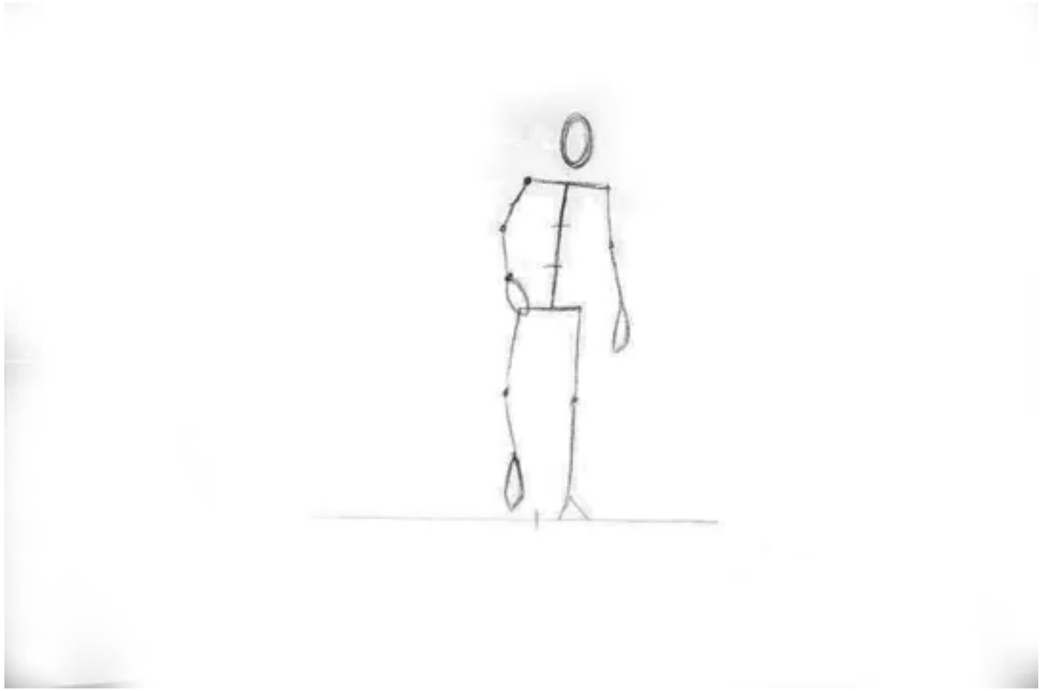


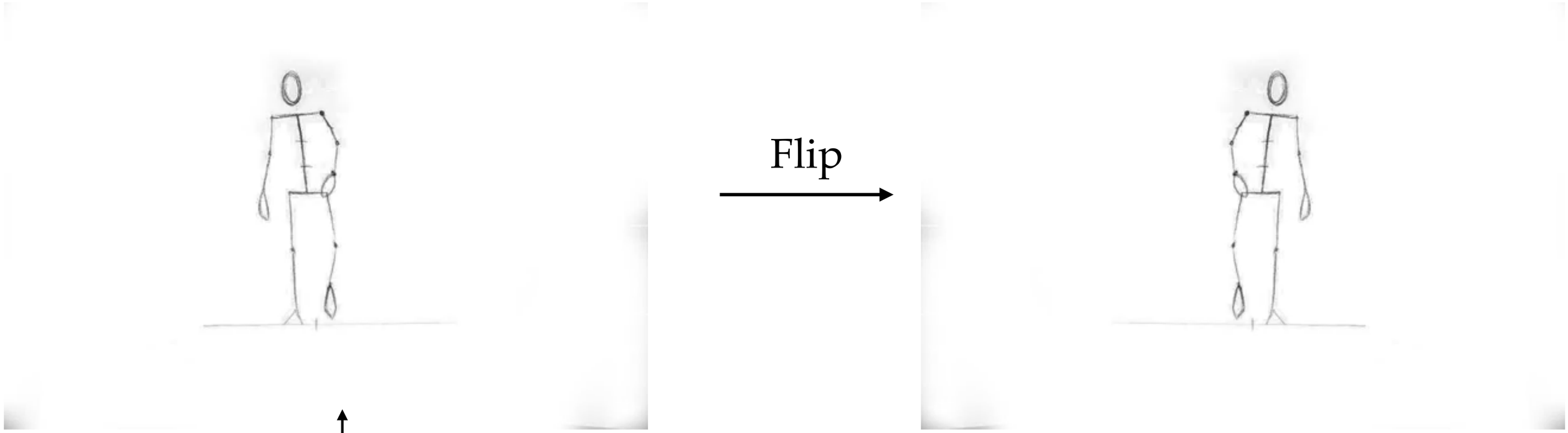
?

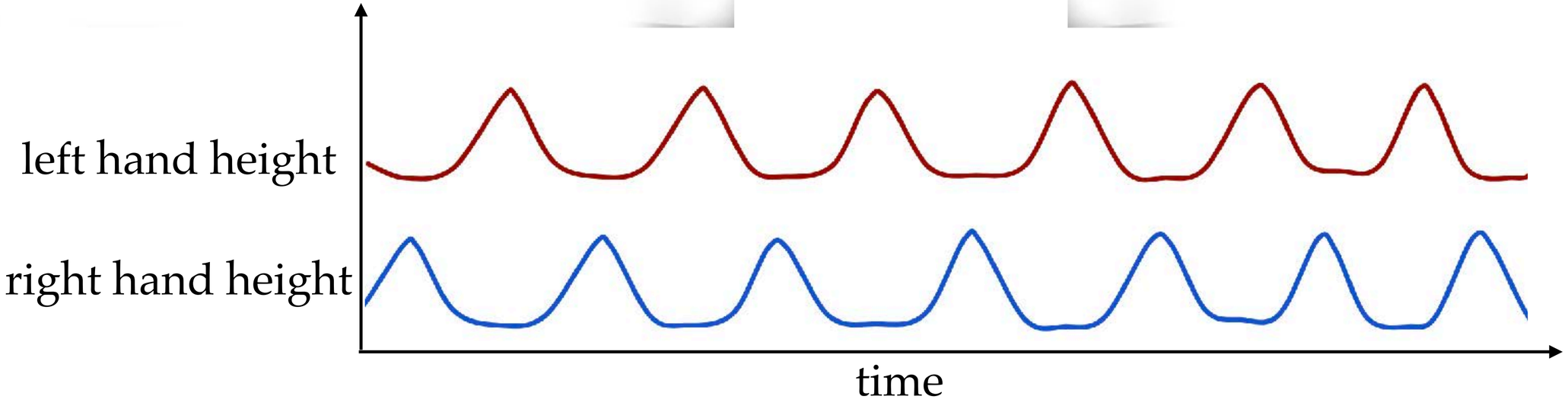


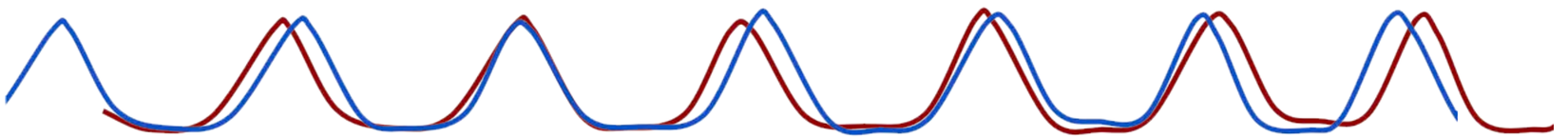


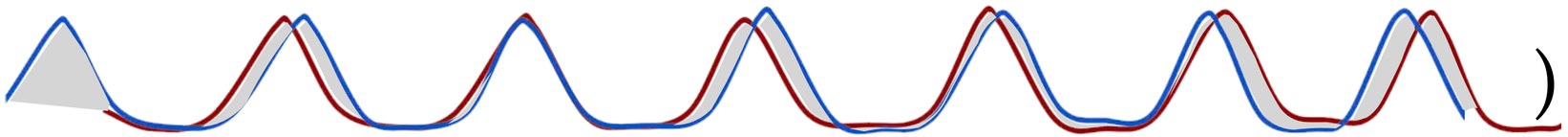
Flip

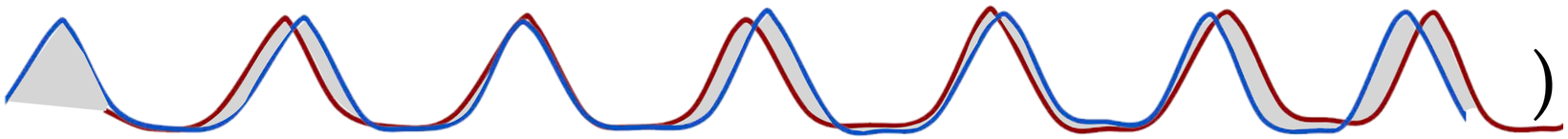


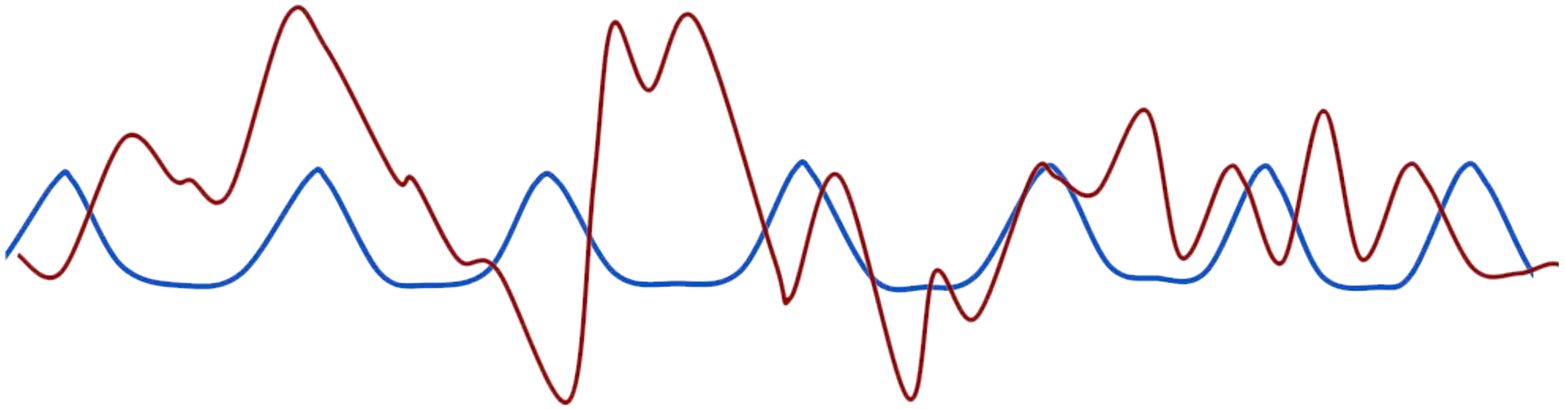

Flip

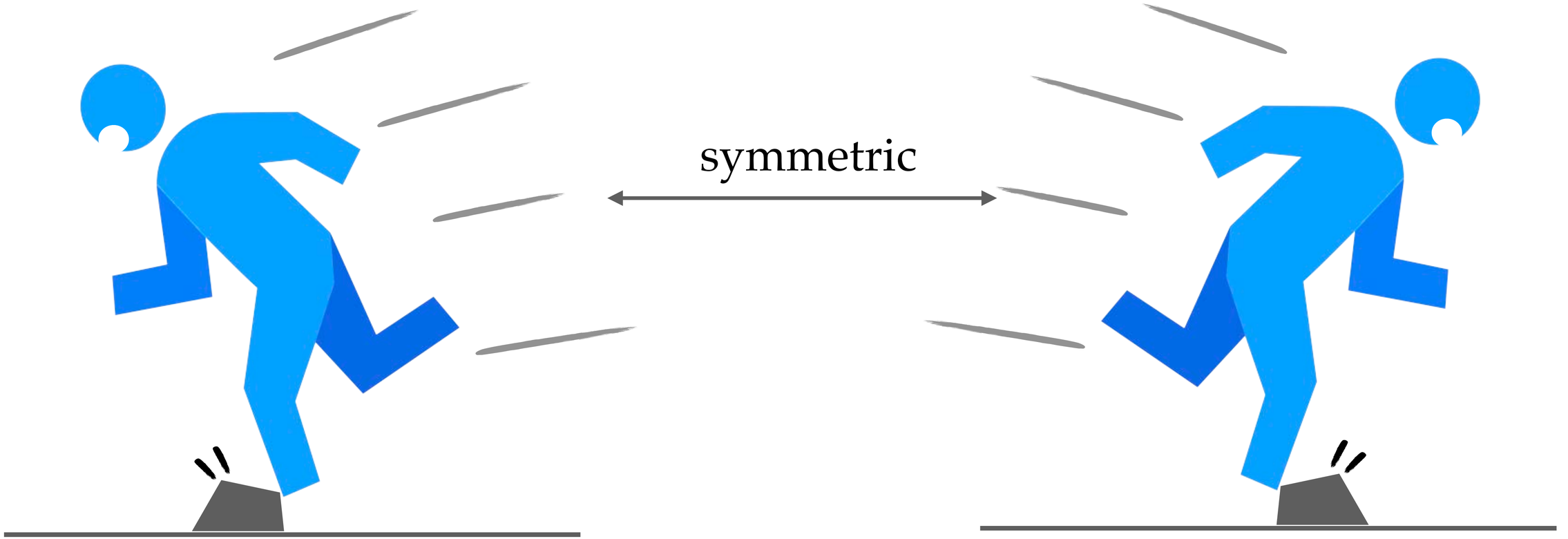

left hand height
right hand height
time

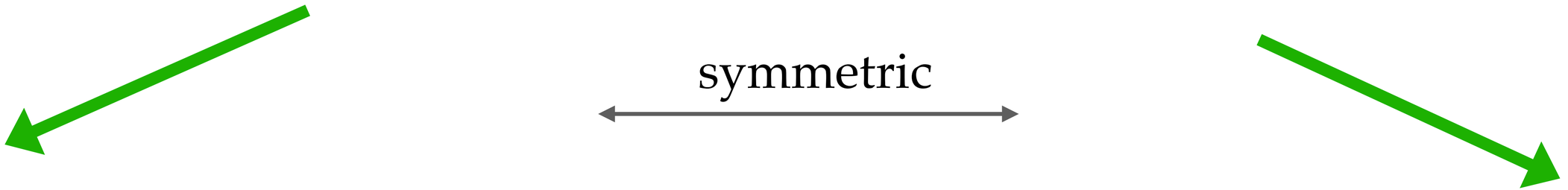


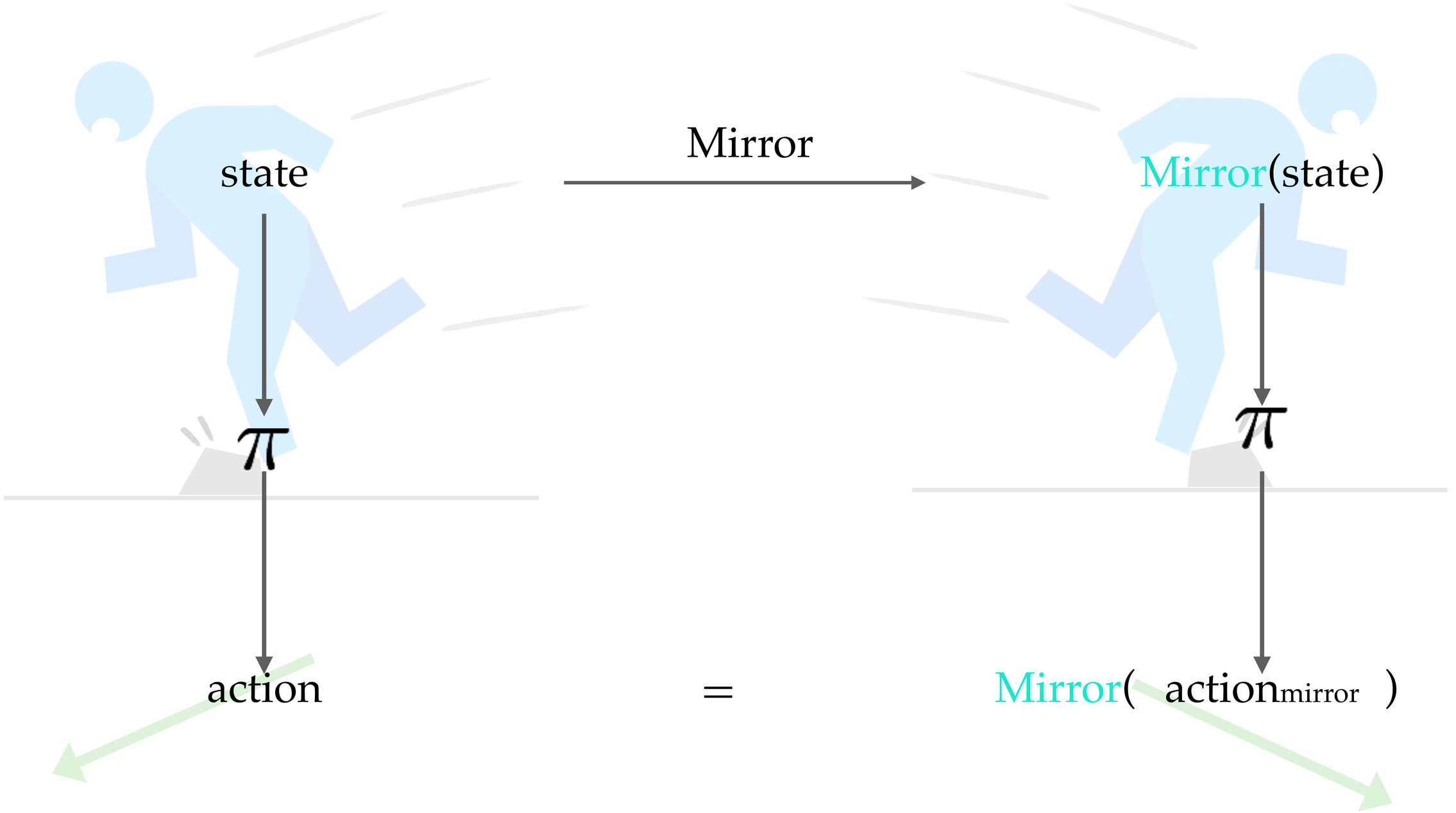
Asymmetry = ()

Reward -= ()

Reward -= ()


symmetric


symmetric


state
Mirror
Mirror(state)
π
π
action
=
Mirror(action mirror)

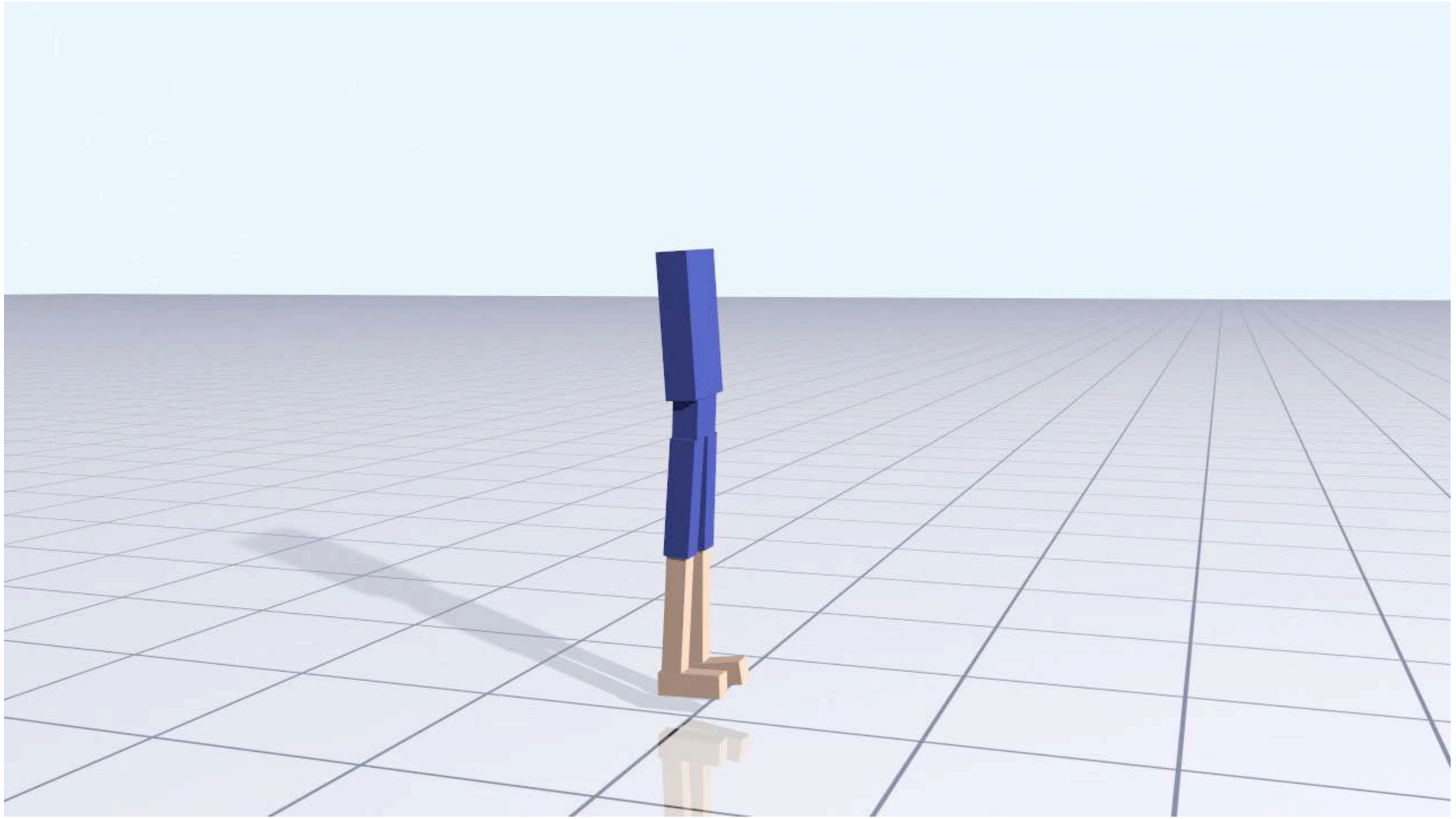
$$\underset{\pi}{\text{minimize:}} \quad L_{RL}$$

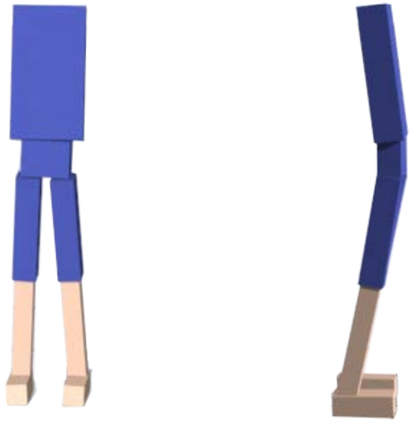
$$\text{subject to:} \quad \text{action} = \text{Mirror}(\ \text{action}_{\text{mirror}}\ )$$

*weights omitted for clarity

$$\underset{\pi}{\text{minimize:}} \qquad L_{RL} + \left\| \text{action} - \text{Mirror}(\ \text{action}_{\text{mirror}}\ ) \right\|_2^2$$

*weights omitted for clarity





Curriculum Learning

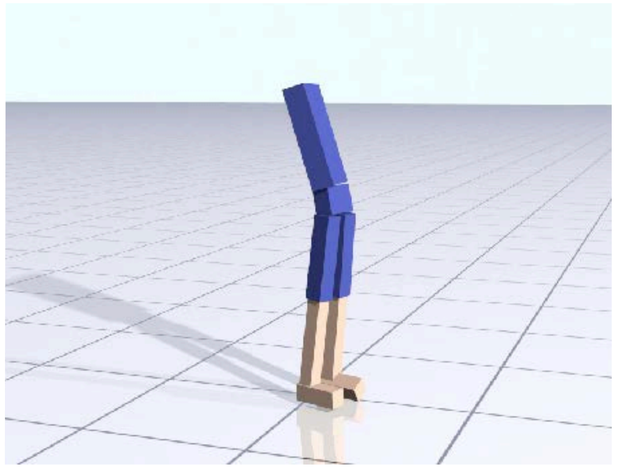
Deep Reinforcement Learning

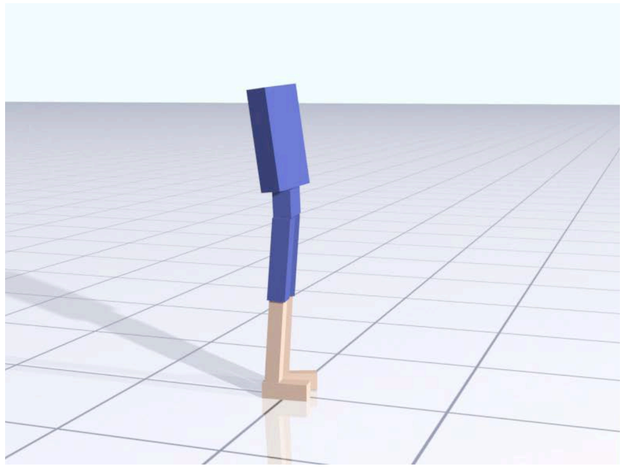
+

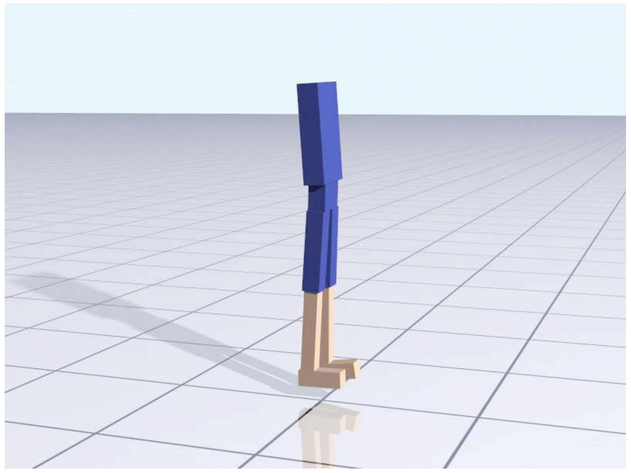
Energy Minimization

+

Gait Symmetry

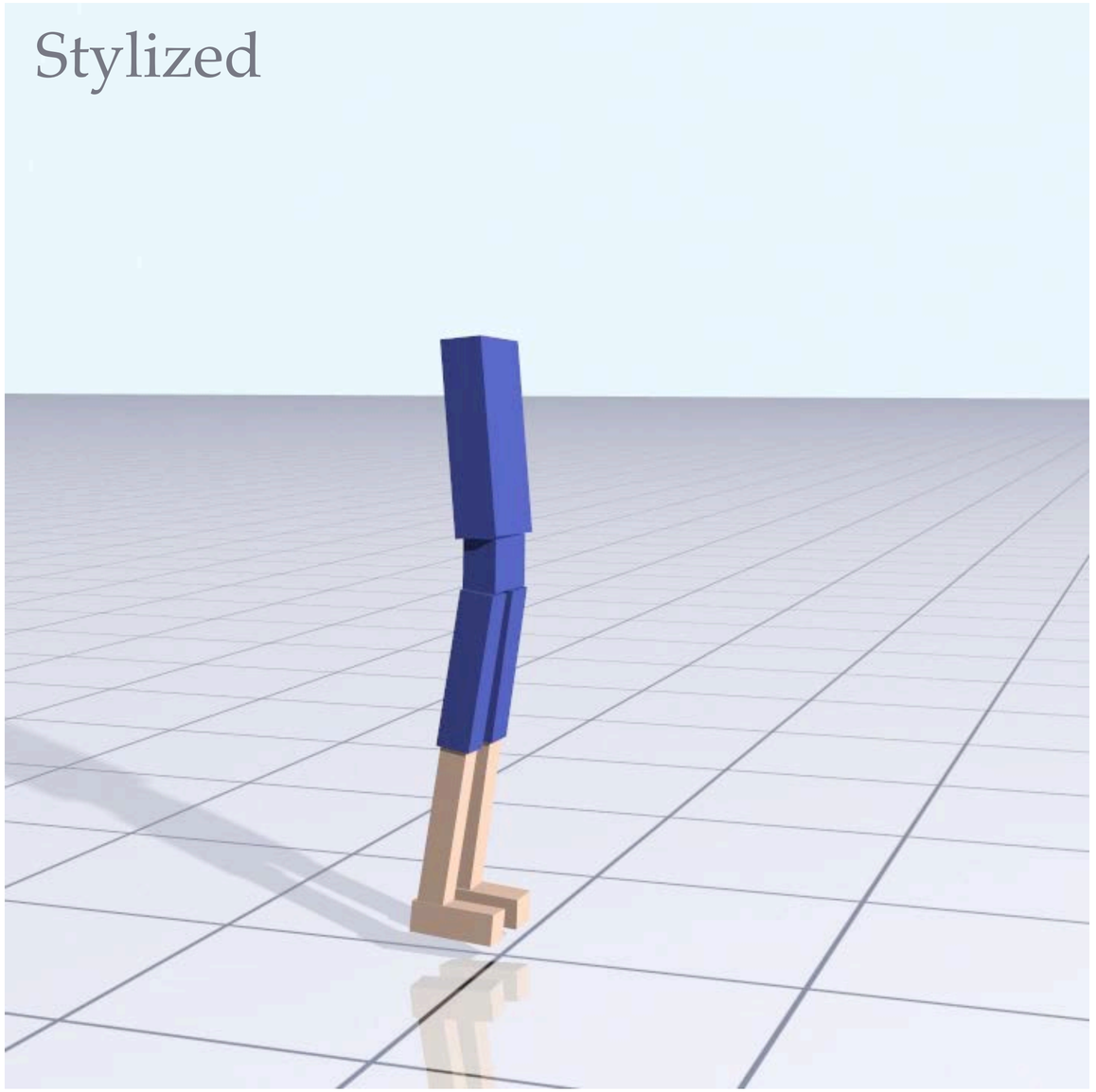




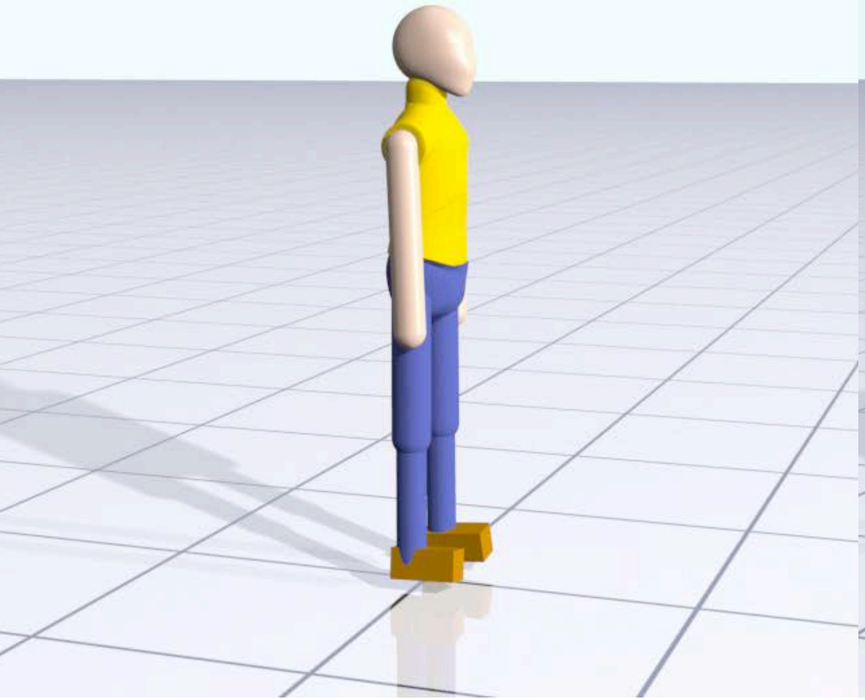


Running

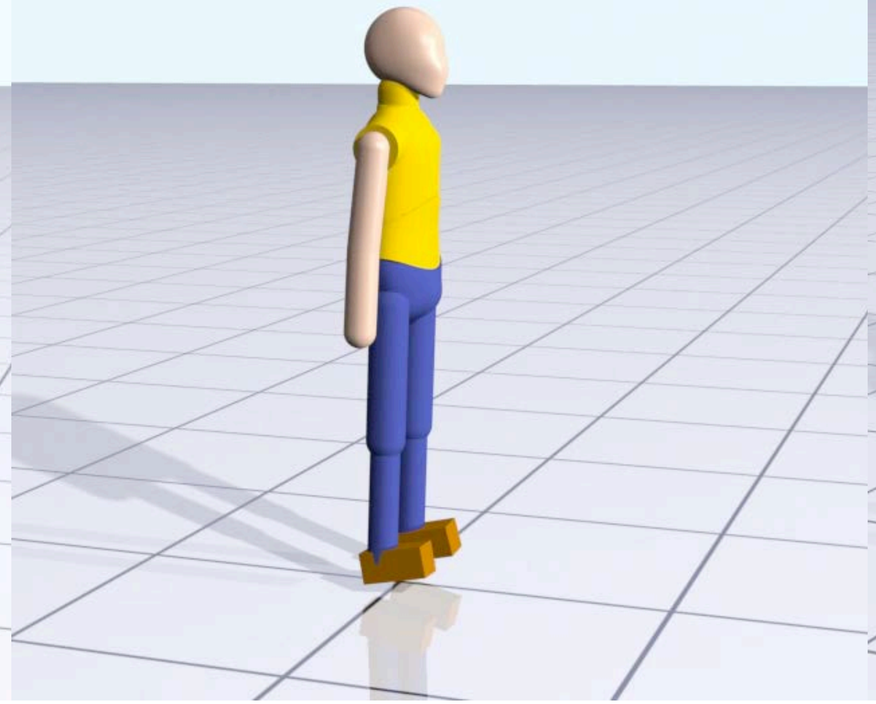
Stylized



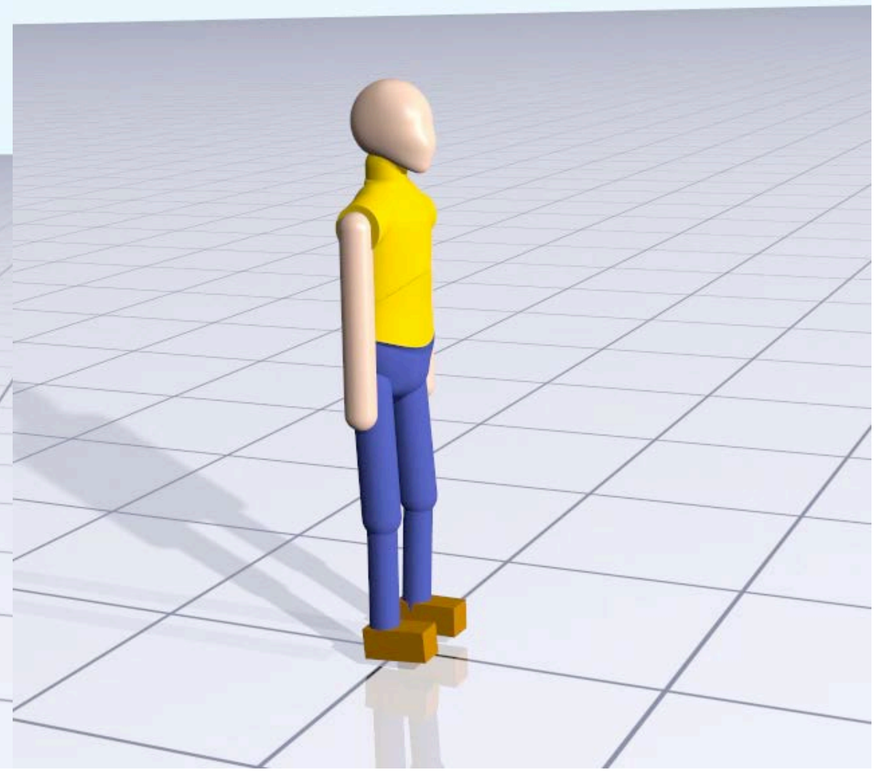
Walking



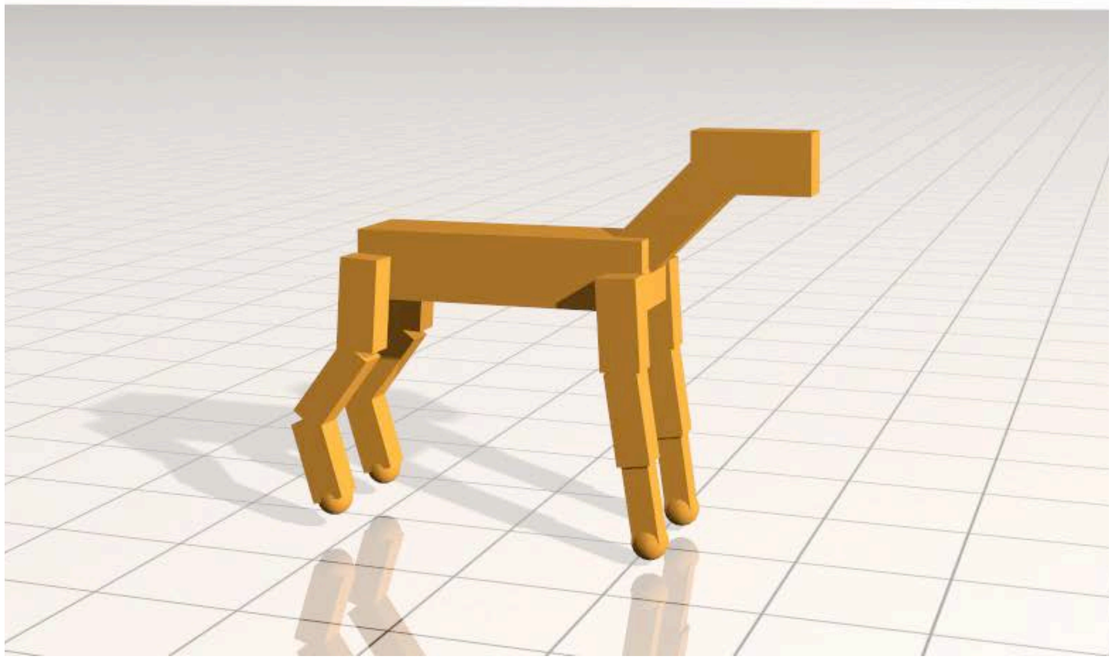
Running



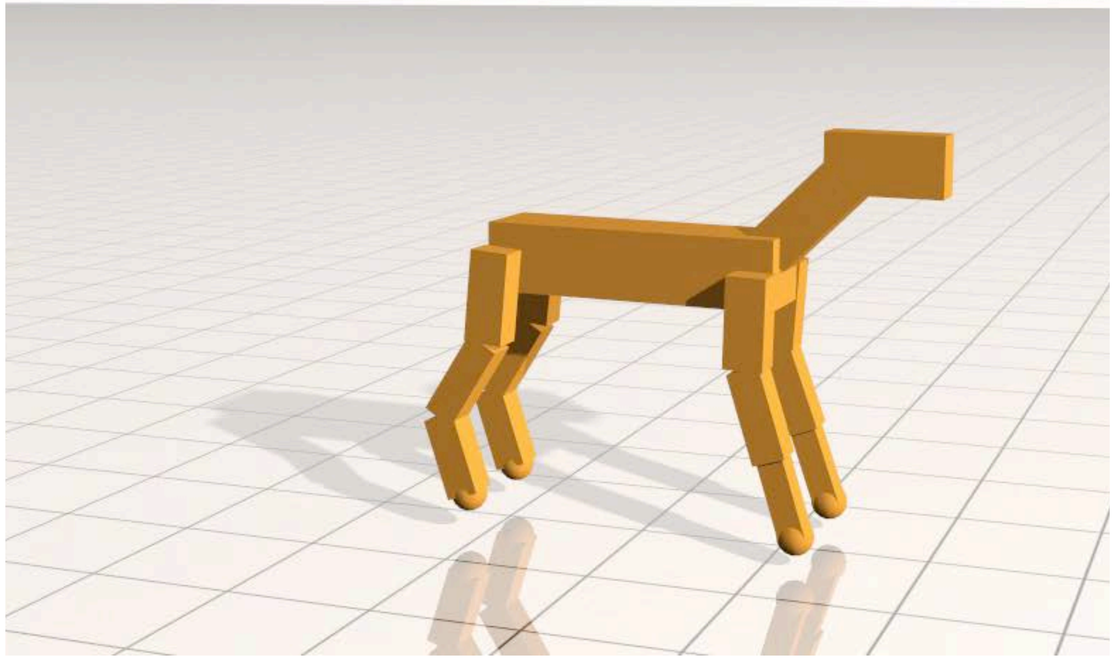
Walking back

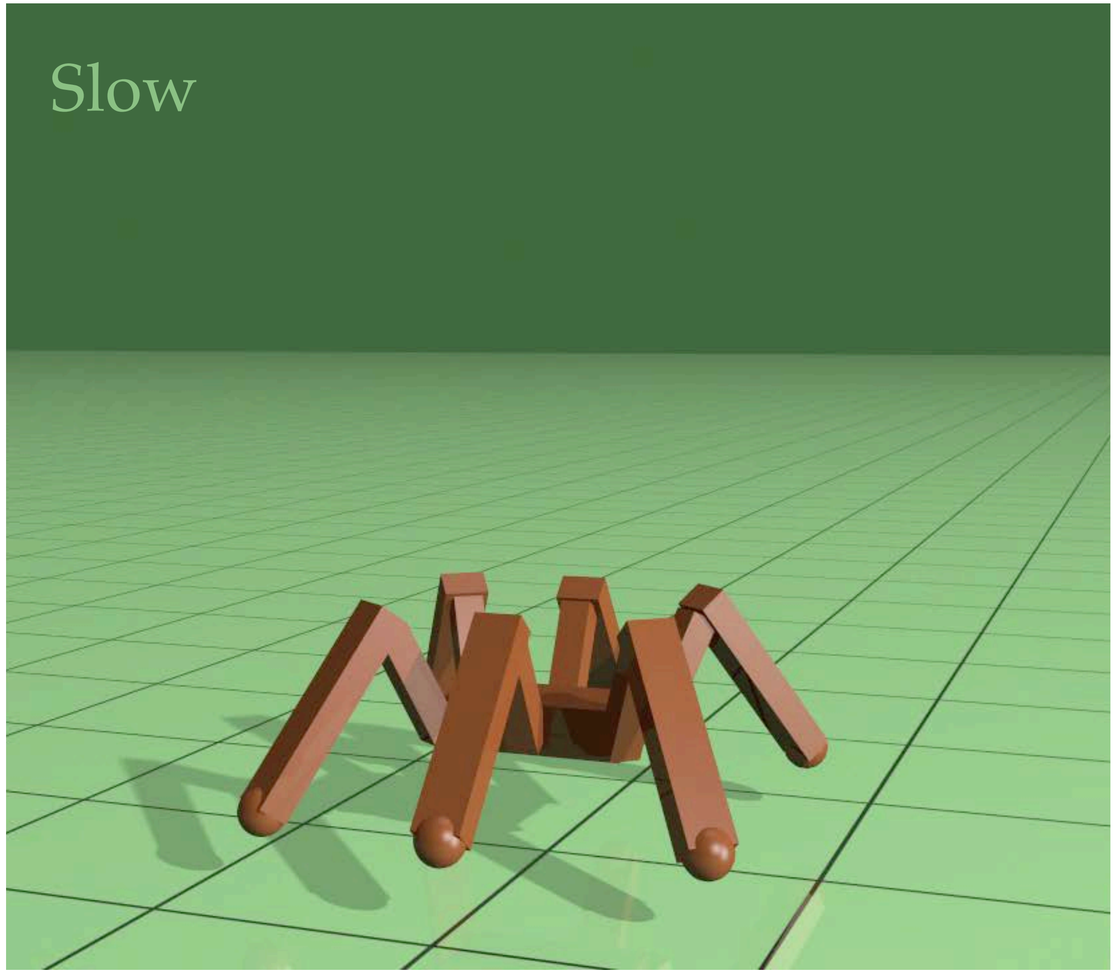


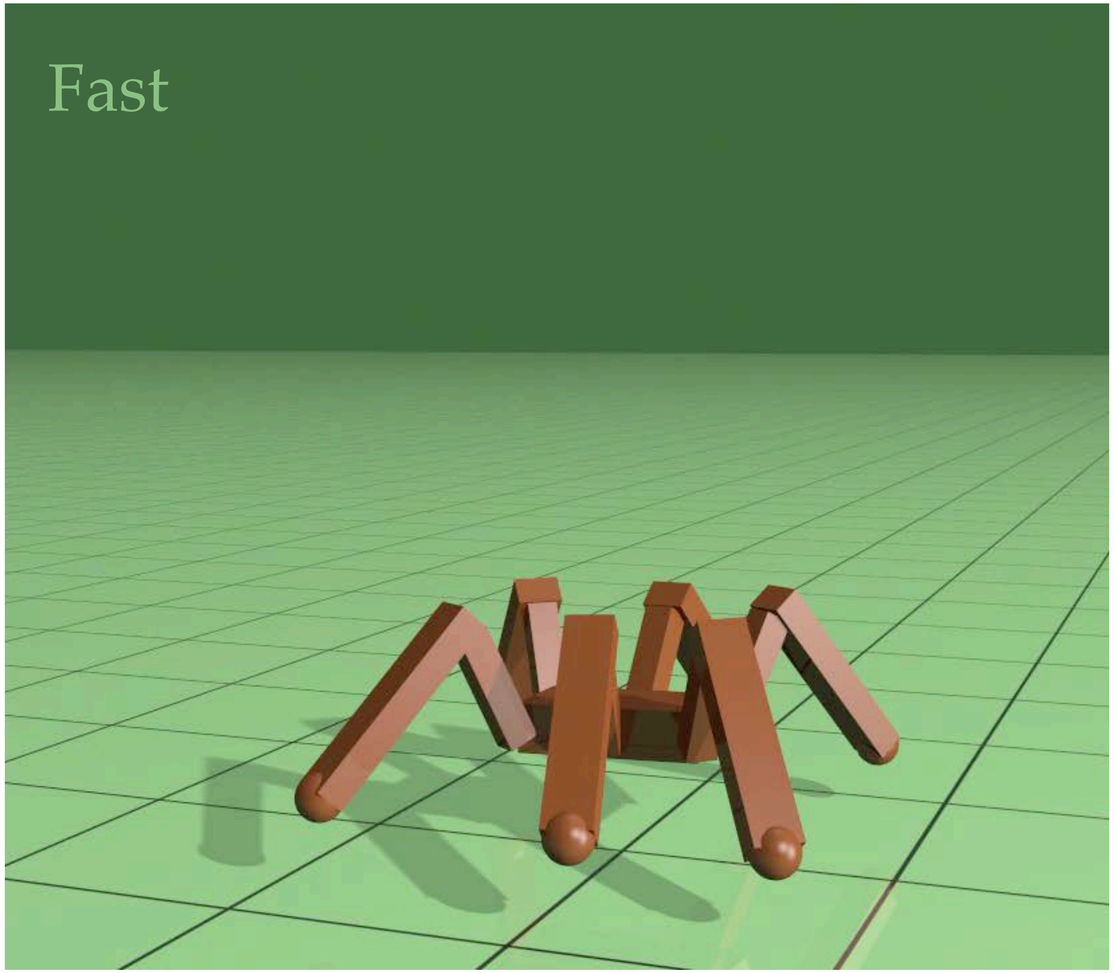
Trotting



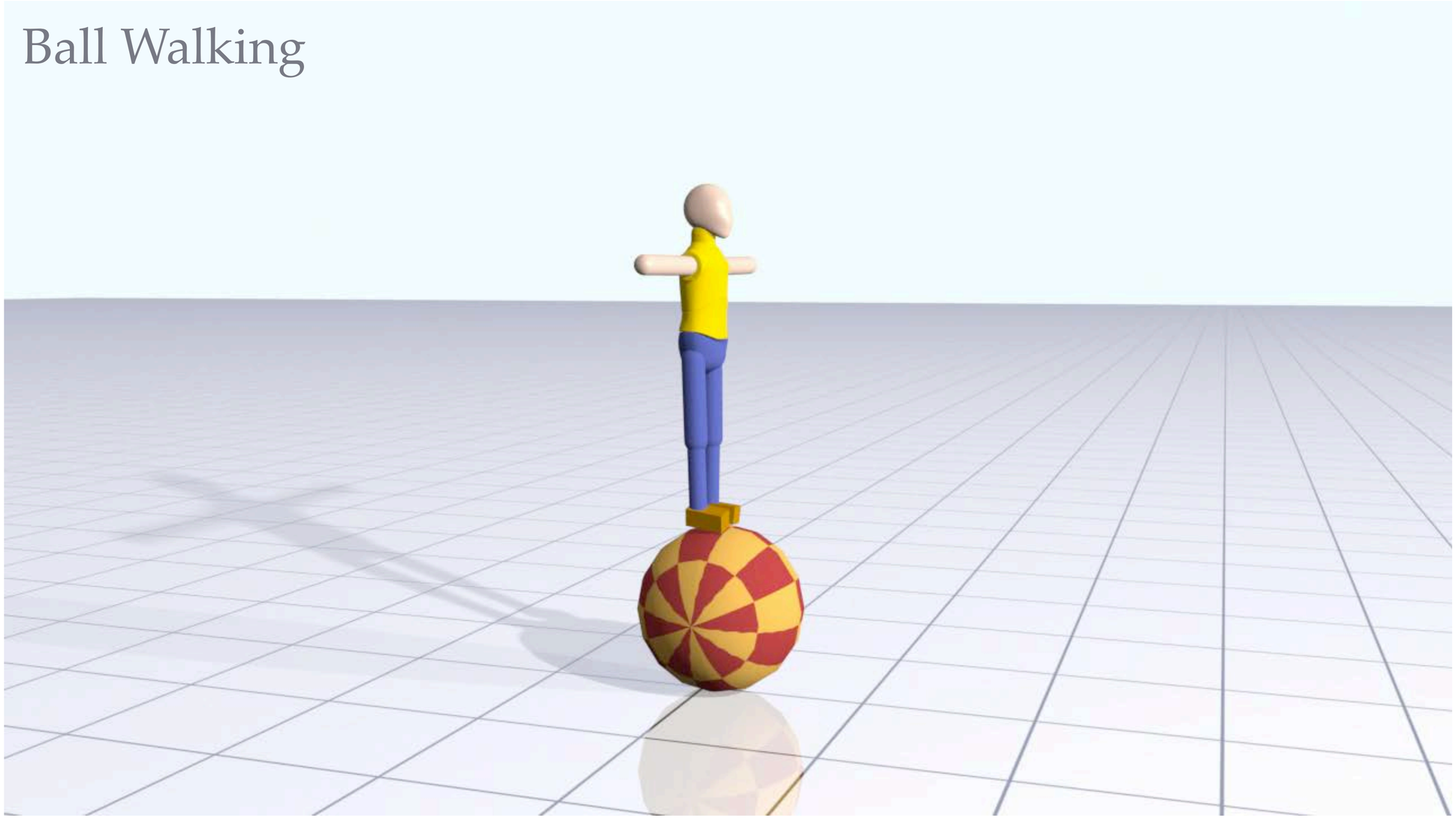
Galloping




Slow

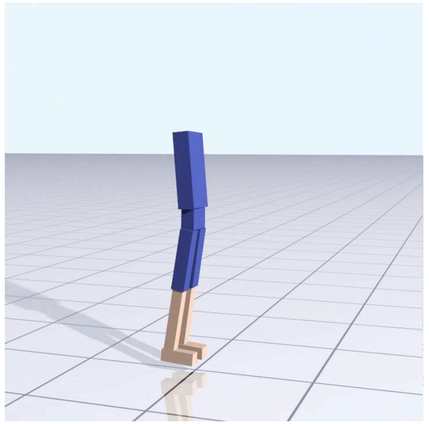
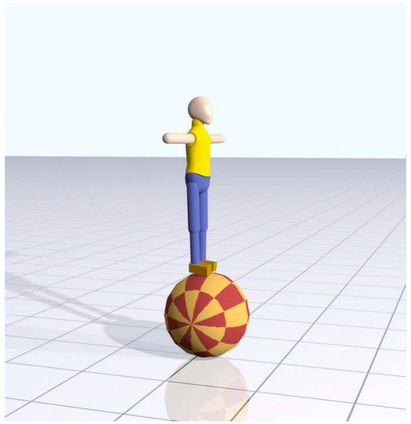
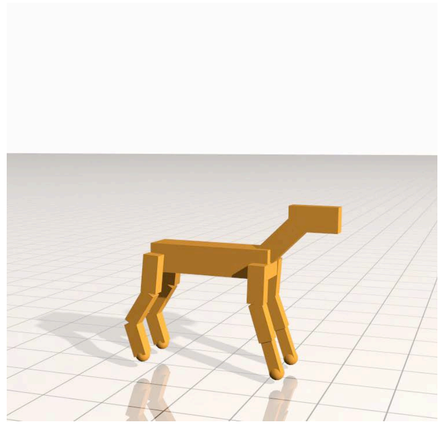

Fast

Ball Walking



What's Next?

- Biomechanics-based models
- Extend to more agile motions such as gymnastics
- Running on real-hardware


Thank you!

- paper link: https://arxiv.org/abs/1801.08093
- code link: https://github.com/VincentYu68/SymmetryCurriculumLocomotion
- my website: wenhaoyu.weebly.com
- email: wenhaoyu@gatech.edu