

SUPPLEMENTARY MATERIAL

10 LLC WALK CYCLES



Fig. 15. LLC walk cycles. **top-to-bottom**: nominal walk, in-place walk, high-knees, straight leg, straight legs, forward lean, sideways lean.

11 LLC REWARD

$$r_L = w_{pose}r_{pose} + w_{vel}r_{vel} + w_{root}r_{root} + w_{com}r_{com} + w_{end}r_{end} + w_{heading}r_{heading}$$

$$r_{pose} = \exp\left(-\sum_i w_i d(\hat{q}_i(t), q_i)^2\right)$$

$$r_{vel} = \exp\left(-\sum_i w_i \|\hat{q}_i(t) - \dot{q}_i\|^2\right)$$

$$r_{root} = \exp\left(-10(\hat{h}_{root} - h_{root})^2\right)$$

$$r_{com} = \exp\left(-\|\hat{v}_{com} - v_{com}\|^2\right)$$

$$r_{end} = \exp\left(-\|\hat{p}_{swing} - p_{swing}\|^2 - \|\hat{p}_{stance} - p_{stance}\|^2\right)$$

$$r_{heading} = 0.5\cos(\hat{\theta}_{root} - \theta_{root}) + 0.5$$

h_{root} represents the height of the root from the ground, v_{com} is the center of mass velocity, p_{swing} and p_{stance} are the positions of the swing and stance foot. The target position for the swing foot $\hat{p}_{swing} = \hat{p}_0$ is provided by the footstep plan, while the target position for the stance foot \hat{p}_{stance} is provided by the reference motion. θ_{root} represents the heading of the root on the horizontal plane, and $\hat{\theta}_{root}$ is the desired heading provided by the footstep plan.

12 HLC SOCCER DRIBBLING REWARD

$$r_H = w_{cv}r_{cv} + w_{cp}r_{cp} + w_{bv}r_{bv} + w_{bp}r_{bp}$$

$$r_{cv} = \exp\left(-\left(\min(0, u_{ball}^T v_{com} - \hat{v}_{com})\right)^2\right)$$

$$r_{cp} = \exp\left(-d_{ball}^2\right)$$

$$r_{bv} = \exp\left(-\left(\min(0, u_{tar}^T v_{ball} - \hat{v}_{ball})\right)^2\right)$$

$$r_{bp} = \exp\left(-d_{tar}^2\right)$$

u_{ball} is a unit vector pointing in the direction from the character to the ball, v_{com} the character's center of mass velocity, and $\hat{v}_{com} = 1m/s$ the desired speed with which the character should move towards the ball. Similarly, u_{tar} represents the unit vector pointing from the ball to the target position, v_{ball} the velocity of the ball, and $\hat{v}_{ball} = 1m/s$ the desired speed for the ball with which to move towards the target. Once the ball is within $0.5m$ of the target and the character is within $2m$ of the ball, then the goal is considered fulfilled and the character receives a constant reward of 1 from all terms, corresponding to the maximum possible reward.

13 LLC STYLIZATION

Learning curves for each stylized LLC is available in Figure 16. Each network is initialized using the LLC trained for the nominal walk. Performance is measured using only the style term c_{style} , measuring the LLC's conformity to the style objectives.

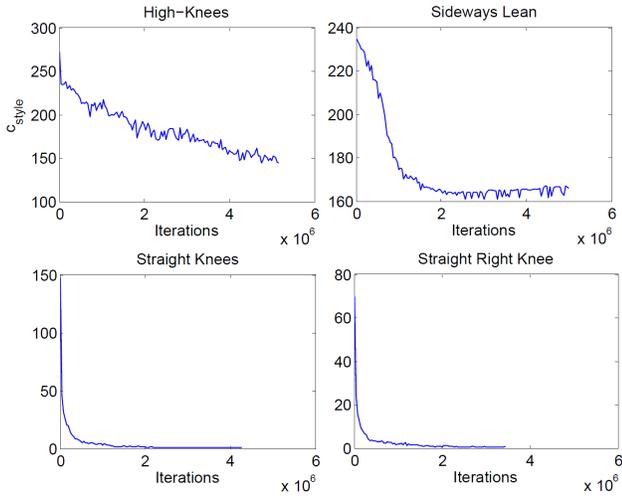


Fig. 16. Learning curves for each stylized LLC.

14 HLC LEARNING CURVES

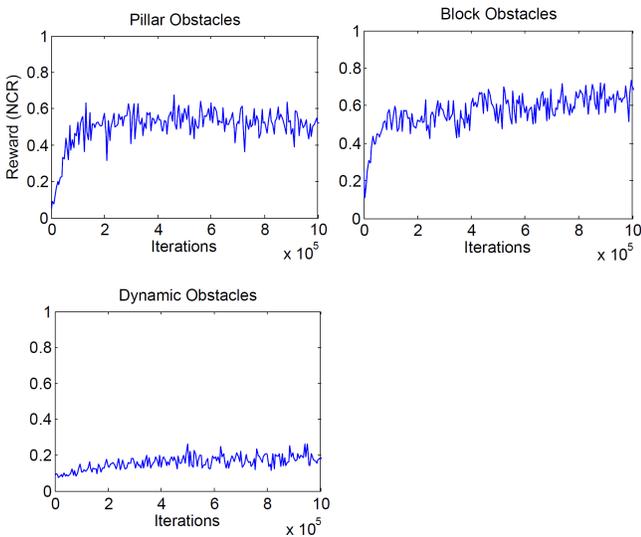


Fig. 17. HLC learning curves.

15 TRANSFER

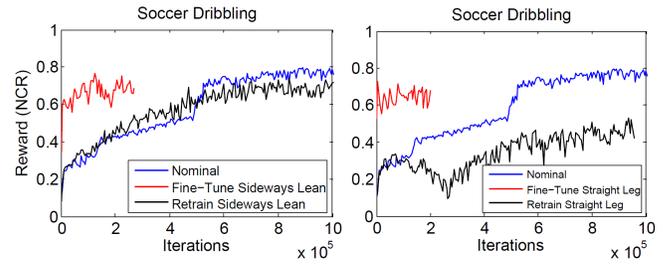


Fig. 18. Learning curves for fine-tuning HLC's for different LLC's, and retraining HLC's from scratch.

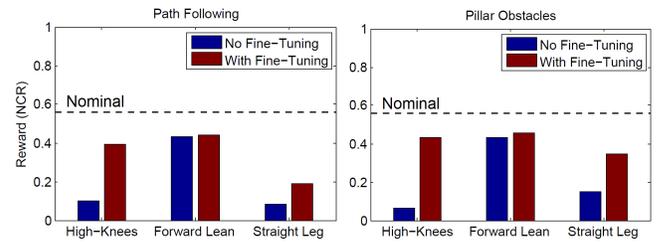


Fig. 19. Performance using different LLC's for path following and pillar obstacles with and without HLC fine-tuning.

Task + LLC	No Fine-Tuning	With Fine-Tuning	Retraining
Soccer + High-Knees	0.19	0.56	0.51
Soccer + Straight Leg	0.51	0.64	0.45
Soccer + Straight Legs	0.16	0.63	-
Soccer + Forward Lean	0.64	0.69	-
Soccer + Sideways Lean	0.41	0.72	0.74
Path + High-Knees	0.10	0.39	-
Path + Forward Lean	0.43	0.44	-
Path + Straight Leg	0.08	0.19	-
Pillars + High-Knees	0.06	0.43	-
Pillars + Forward Lean	0.43	0.45	-
Pillars + Straight Leg	0.15	0.35	-

Table 3. Performance (NCR) of different combinations of LLC's and HLC's. **No Fine-Tuning**: directly using the HLC's trained for the nominal LLC. **With Fine-Tuning**: HLC's fine-tuned using the nominal HLC's as initialization. **Retraining**: HLC's are retrained from random initialization for each task and LLC.