Learning a Walk Behavior Utilizing Toes From Scratch

Jens Fischer, Klaus Dorer

Hochschule Offenburg, Germany, contact: klaus.dorer@hs-offenburg.de

To our knowledge, there is currently no team in the 3D simulation league that actively uses NaoToe's toe joints while walking. In theory, using the toes should allow for a faster, more human-like walk with longer steps. To test this, we adapted a machine learning approach we already used previously to learn kicking while walking [1]. This was relatively straightforward because the approach is model-free and does not make assumptions about the type of behavior that is learned. It uses plain genetic algorithms to learn a movement that is partitioned into four phases of variable length, each phase containing the target speed and angle for each active joint. The movement then repesents a single step, which is mirrored to the other leg and repeated in a loop to create a walking motion. Fitness is defined as distance traveled on the X axis, with penalties for not walking straight and falling.

We were able to achieve speed increases of up to 30% when walking forward, with an average speed of 1.3 m/s instead of our previously fastest 1.0 m/s. With 0.43 m instead of 0.13 m, the step length is also increased considerably (see Figure 1). With minor adjustments to the learning environment, we were also able to learn behaviors for walking backwards and turning. However, the learned walk is not used during the competition due to being roughly 25% less stable. Another non-trivial issue is "morphing" back to other behaviors while running. So far, the most promising solution seems to be to use the same machine learning approach to train a "morph movement", whose 4 phases are executed just once without the usual mirroring and repetition. This results in a quick "jump" that stabilizes the agent and lets him continue with the slower walk of the old walk engine, butb only works in about 50% of cases.

FC Portugal has also learned walk behaviors from scratch with a model-free machine learning approach [2]. However, there are significant differences between the two approaches. While we use genetic algorithms, FC Portugal uses a deep reinforcement learning algorithm called proximal policy optimization (PPO). So far they have not utilized NaoToe / toe joints. Despite this, their walk behavior is still significantly faster with over 2 m/s and more stable.

References

1. Klaus Dorer. *Learning to Use Toes in a Humanoid Robot.* In Proceedings of the RoboCup Symposium, Nagoya, Japan, 2017.
2. Miguel Abreu, Nuno Lau, Armando Sousa, and Luis Paulo Reis. *Learning low level skills from scratch for humanoid robot soccer using deep reinforcement learning.* In Proceedings of the 2019 International Conference on Autonomous Robot Systems and Competitions (ICARSC).

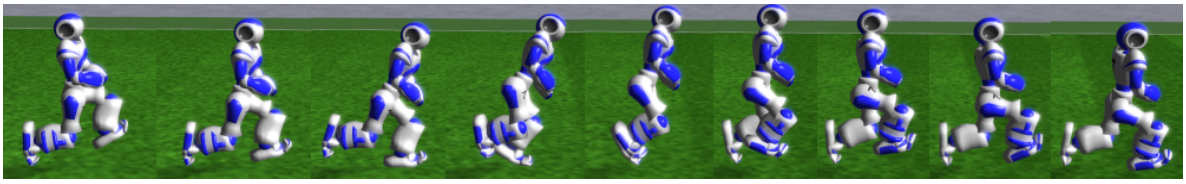

Fig. 1. The learned forward walk movement with toe joint usage.