Bipedal Locomotion as a Metastable Markov Decision Process

Cenk Oguz Saglam, Katie Byl
University of California, Santa Barbara
{saglam,katiebyl}@ece.ucsb.edu

***Abstract*—We present and explain tools we find useful for modeling and optimizing rough terrain walking. Our aim is to give an intuitive overview, toward encouraging use by others.**

I. Introduction

It is hard to predict long-term dynamics for hybrid, stochastic, nonlinear dynamic systems - such as walking robots. Given realistic noise and actuator limits, guarantees of global stability often cannot be given. Rather than abandoning all hope, we have proposed modeling very good walking as a metastable process, with reliability quantified by the mean first-passage time (MFPT) between falls [1]. Here, we briefly overview our current methods and rationale for MFPT estimation; i.e., we model step-to-step dynamics as a Markov Decision Process, and the resulting dynamics for metastable walking are typically dominated by a single (slow) discrete-time pole.

II. Meshing, Switching, and Metastability

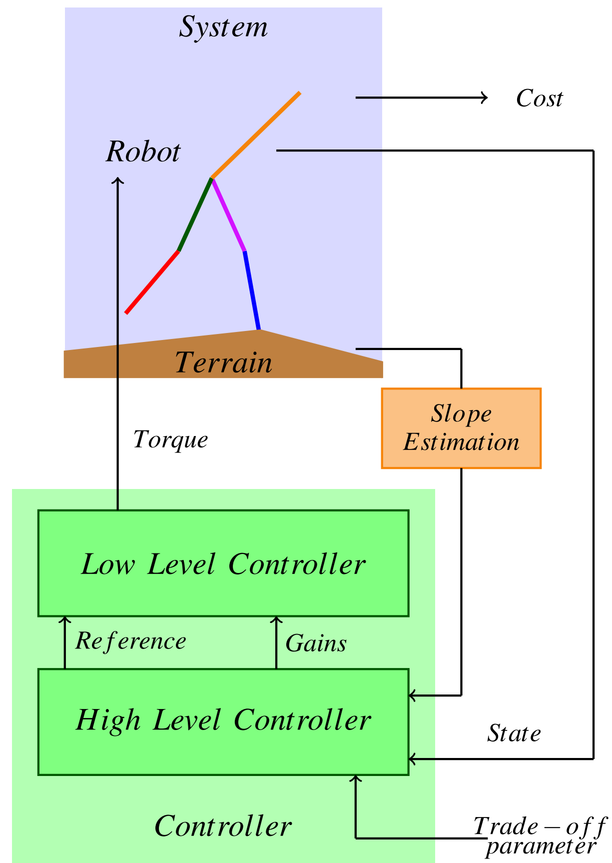
Monte Carlo simulations are not practical to estimate MFPT for metastable systems, motivating our Markov chain modeling. For our approach, this requires the existence of low-level controllers. For piecewise-sloped ground, the next state of the robot, $x[n+1]$, is a function of the current state $x[n]$, the slope ahead $\gamma[n]$, and the controller used $\zeta[n]$, i.e.,

$$x[n+1] = h(x[n], \gamma[n], \zeta[n]). \tag{1}$$

To mesh, we define a Poincaré section immediately before each foot impact. A limit cycle in the full dynamics is a fixed point in this section. Starting with a fixed point and an absorbing failure state (i.e., two points) as our initial mesh, we simulate for each controller/slope combination and add resulting states if a distance metric to all existing nodes is large enough, repeating until no new point is added, i.e., until the mesh covers the reachable part of the state space. Using the mesh, any given control policy that switches on a step-to-step basis between our controllers can be represented as a stochastic transition matrix, and for metastable systems

$$MFPT = \frac{1}{1-\lambda_2}, \tag{2}$$

where λ_2 is the second-largest eigenvalue of this matrix. Given the pre-calculated dynamics on the mesh, we can find optimal policies (based on robot state and/or partial terrain info) and can quantify effects of (for example) bad sensing, bad terrain, and changes to the available low level controllers [2, 3, 4].



Acknowledgments

This work is supported by the Institute for Collaborative Biotechnologies (ICB) through grant W911NF-09-0001 from the U.S. Army Research Office. Contents do not necessarily reflect the position or policy of the Government.

References

[1] K. Byl and R. Tedrake. Metastable walking machines. *Int J Robotics Research*, 28:1040–1064, Aug 2009.
[2] C. O. Saglam and K. Byl. Stability and gait transition of the five-link biped on stochastically rough terrain using a discrete set of sliding mode controllers. In *ICRA*, 2013.
[3] C. O. Saglam and K. Byl. Switching policies for metastable walking. In *CDC*, 2013.
[4] C. O. Saglam and K. Byl. Robust policies via meshing for metastable rough terrain walking. In *RSS*, 2014.