Learning Locomotion Primitives from Contextual Bayesian Optimization

Brian Yang
UC Berkeley

Grant Wang
UC Berkeley

Roberto Calandra
UC Berkeley

Daniel S. Contreras
UC Berkeley

Sergey Levine
UC Berkeley

Kristofer S. J. Pister
UC Berkeley

Abstract

The design of gaits for robot locomotion can be a daunting process that can be addressed by data-driven gait optimization. In this paper, we propose a novel approach to efficiently learn a wide range of locomotion tasks with walking robots. This approach formalizes locomotion as a contextual Bayesian optimization task to collect data, and subsequently uses that data to learn multi-objective locomotion primitives that can be used for planning. As a proof-of-concept, we demonstrate on a simulated micro-hexapod that without any prior knowledge about the robot used (e.g., dynamics model), our approach is capable of learning locomotion primitives within 250 trials and then use them to successfully navigate through a maze.

1 INTRODUCTION

Substantial progress has been made in developing fully autonomous microrobots [17, 22]. However, the design and implementation of gaits for enabling locomotion at the sub-centimeter scale remains a non-trivial task. Our primary contribution is to introduce a novel approach that allows for efficient learning of gaits and motor primitives from scratch without prior knowledge. This is accomplished by collecting data on various motor primitives using contextual Bayesian optimization and using those evaluations to reformulate the problem into a multi-objective optimization task, giving us a model that can map any set of parameters to a predicted trajectory. Using this model, we are able to optimize our parameters on various trajectories for subsequent use in path planning. To evaluate our approach, we used a simulated hexapod microrobot modeled after a recently developed microrobot [6]. We first validated existing techniques on a set of progressively more difficult tasks: learning single-objective, contextual, and multi-objective gaits. Following, we evaluated our approach by learning motor primitives from 250 trials, and used them to successfully navigate a maze.

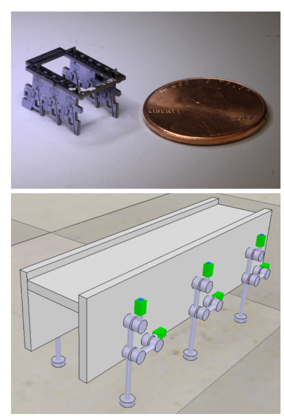


Figure 1: The six-legged micro walker considered in our study (top) and its simulation (bottom).

2 RELATED WORK

For controllers with many parameters, manually tuning controller parameters can require both an immense amount of domain expertise and time. As such, automatic gait optimization is an important research field which has been studied with a wide variety of approaches in both the single-objective [19, 5, 12, 10, 20, 13, 14, 2] and multi-objective setting [4, 13, 14, 21]. A more data-efficient approach used in the past to learn gaits for snake and bipedal robots is Bayesian optimization [10, 20, 2]. Bayesian optimization has also been applied to contextual policy search in the

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

context of robot manipulation [11]. Our contribution builds off of this work by applying and extending the contextual framework to learning movement trajectories and path planning. Another extension of Bayesian optimization related to our work is Multi-objective Bayesian optimization, which has also been previously applied in the context of robotic locomotion [21]. Our main contribution demonstrates an entirely novel application of multi-objective optimization from past work by using a multi-objective model to learn over an area of possible trajectories for path planning.

3 BACKGROUND

Central Pattern Generators Central pattern generators (CPGs) are neural circuits found in nearly all vertebrates, producing periodic outputs without sensory input and can reproduce a variety of gaits [24]. Benefits of using a CPG controller are lack of computational intensity and drastic reduction in number of optimization parameters $\boldsymbol{\theta}_i$. The 4 parameters we consider during optimization are $\boldsymbol{\theta} = [\omega, X, R_l, R_r]$ where ω is frequency of the oscillators and X is phase difference between each of the vertical-horizontal oscillator pairs. In order to allow for directional control, R_l and R_r are the amplitudes of the left and right side oscillators respectively. We built our controller using a network of 12 coupled phase oscillators (1 per motor) and implemented four gaits – tripod, ripple, wave, and four-two. For more information about gaits we refer the reader to [3].

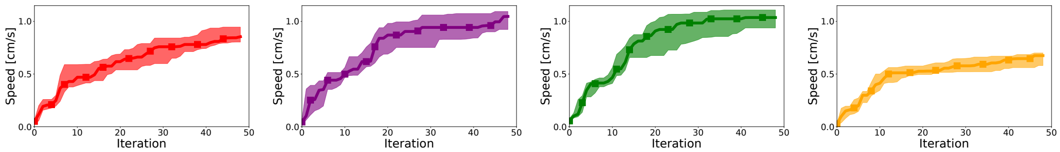
Bayesian Optimization To automate the CPG network's parameters tuning, we use Bayesian optimization (BO) [9, 7, 2]. We formulate the tuning of the CPG parameters as the optimization $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} f(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ are the parameters to be optimized w.r.t. the chosen objective function f (e.g., walking speed, which we investigate in Section 5). BO learns a model $\tilde{f} : \boldsymbol{\theta} \rightarrow f(\boldsymbol{\theta})$ at each iteration, and the model $\tilde{f}$ is then optimized using a Gaussian Process [16]. The resulting set of parameters $\boldsymbol{\theta}^*$ from the optimized model is then evaluated on the real system and together with $f(\boldsymbol{\theta}^*)$ is added to the data set. For more information regarding BO, we refer the readers to [7, 18].

Multi-objective Bayesian Optimization A special case of the optimization task is multi-objective optimization (MOO) [1]. Often in robotics, multiple conflicting objectives need to be optimized simultaneously, causing design trade-offs (e.g., walking speed vs energy efficiency which we investigate in Section 5). When multiple objectives are accounted for, there is no single optimum solution, but a set of Pareto optimal solutions [15], known as the Pareto front (PF). MOO often requires a significant number of experiments, so we turn to ParEGO, a multi-objective Bayesian optimization algorithm that addresses this issue [8]. For more information about multi-objective Bayesian optimization we refer the reader to [23].

Contextual Bayesian Optimization Another case of the optimization task is contextual optimization, where we assume there are multiple correlated, but different tasks identified by context variable $\boldsymbol{c}$. An example (which we investigate in Section 5) is walking on inclines, where the context variable is the slope angle. We can formalize this as $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} f(\boldsymbol{\theta}, \boldsymbol{c})$, where for each context $\boldsymbol{c}$, a set of parameters $\boldsymbol{\theta}^*$ exists. We can exploit the correlation between the tasks to generalize and quickly learn to solve new contexts. Here we consider contextual Bayesian optimization (cBO) [11] which extends the BO framework from Section 3. Utilizing cBO, we will show in Section 5 that our microrobot can learn to walk uphill and curve.

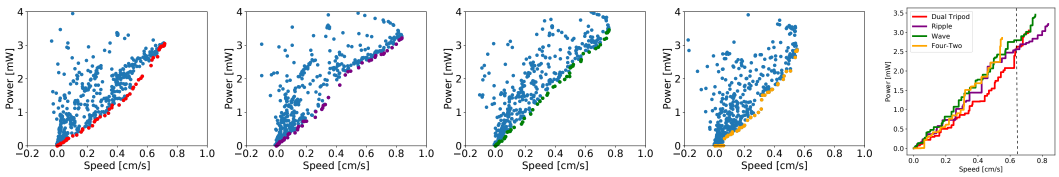
4 LEARNING LOCOMOTION PRIMITIVES FOR PATH PLANNING

We now present our novel approach to learn motor primitives for path planning by re-using the evaluations collected using cBO to convert the task into a MOO problem. Let us consider a cBO task where we want to optimize the parameters $\boldsymbol{\theta}$ to reach different target positions $\boldsymbol{c} = [\Delta x_{\text{des}}, \Delta y_{\text{des}}]$ (this setting is evaluated in Section 5). The objective function is defined as the Euclidean distance $f = \sqrt{(\Delta x_{\text{des}} - \Delta x_{\text{obs}})^2 + (\Delta y_{\text{des}} - \Delta y_{\text{obs}})^2}$ where $\Delta x_{\text{obs}}, \Delta y_{\text{obs}}$ are the actual positions measured after evaluating a set of parameters. The cBO model maps $\tilde{f} : [\boldsymbol{\theta}, \Delta x_{\text{des}}, \Delta y_{\text{des}}] \rightarrow f(\boldsymbol{\theta})$. However, to compute f it needs to measure $\Delta x_{\text{obs}}, \Delta y_{\text{obs}}$, effectively generating data of the form $[\boldsymbol{\theta}, \Delta x_{\text{des}}, \Delta y_{\text{des}}] \rightarrow [\Delta x_{\text{obs}}, \Delta y_{\text{obs}}, f(\boldsymbol{\theta})]$. We can re-use the data generated from cBO to learn a motor primitive model in the form $g : \boldsymbol{\theta} \rightarrow [\Delta x_{\text{obs}}, \Delta y_{\text{obs}}]$. The purpose of this learned model g is to provide an estimate of the final displacement obtained for a set of parameters independently from the optimization process that generated it. Once such a model is learned, we can use it to compute parameters that lead to a desired displacement $\Delta x^*_{\text{obs}}, \Delta y^*_{\text{obs}}$ by optimizing the parameters w.r.t. the



(a) Dual Tripod (b) Ripple (c) Wave (d) Four-Two

Figure 2: Learning curve for the four gaits (median and 65th percentile). We can see how, for all the gaits, BO learns to walk from scratch within 50 iterations. After the optimization, Wave and Ripple are the fastest gaits at $\sim 1\,\mathrm{cm/s}$.



(a) Dual Tripod (b) Ripple (c) Wave (d) Four-Two (e) PF Comparison

Figure 3: Performance measured for the four gaits, and the corresponding PFs. ParEGO is able to quickly explore the PF for each of our four gaits.

output of the model $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} z(g(\boldsymbol{\theta}))$, where z is a scalarization function of our choice (e.g., the Euclidean distance). These paraemters are optimized over a series of multiple displacements to obtain a path planning optimization. In Section 5, we employ a simple shooting method optimization which randomly samples multiple candidate parameters and selects the best outcome.

5 EXPERIMENTAL RESULTS

Here we discuss the performance of our simulated microrobot on various locomotive tasks. Videos of tasks are available at `https://sites.google.com/site/learninglocomotorprimitives/`.

Learning to Walk Straight We optimized the four gaits considered using as our objective function the robot's walking speed (measured as distance traveled after $1\,\mathrm{s}$). The optimization used the 4 parameters outlined in Section 3 and was repeated 20 times for each gait, as seen in Figure 2, Results show that the optimizer learned to walk from scratch within 50 iterations. Moreover, we note that the optimized wave and ripple are the fastest gaits at $\sim 1\,\mathrm{cm/s}$.

Multi-objective Gait Optimization We now consider a multi-objective optimization setting and compare the different gaits w.r.t. the microrobot's walking speed and motor energy consumption. We optimized the four gaits using multi-objective Bayesian optimization with the same 4 parameters as before on a budget of 50 iterations. In Figure 3 we can see the Pareto fronts obtained for the different gaits. From these results, we observe that the tripod gait dominates other gaits in speed $< 0.6\,\mathrm{cm/s}$, while Ripple dominates when speed is $> 0.6\,\mathrm{cm/s}$, hence a clear indication of which gait is preferable under different circumstances.

Discovering New Gaits with Multi-objective Optimization We also tested multi-objective optimization on the walker without using predefined gaits. The resulting multi-objective optimization task has 8 parameters (frequency, phase difference between horizontal and vertical motors, and the six gait coupling parameters). In Figure 4 we can see the Pareto front for the resulting gaits. Even while penalizing curved paths, the fastest discovered gait outperformed Ripple (the fastest nature-inspired gait we found) by almost 50%.

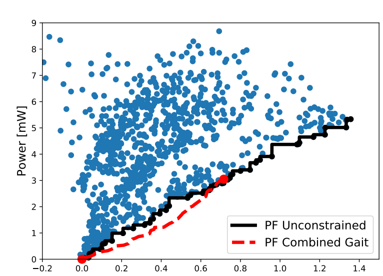


Figure 4: PF of the unrestrained gait optimization versus the best performance of the four nature-inspired gaits.

Learning to Walk on Inclined Surfaces We now consider cBO, specifically gait optimization for walking on inclined terrain, where the angle of the inclination is the context. Sampling randomly over a continuous interval of

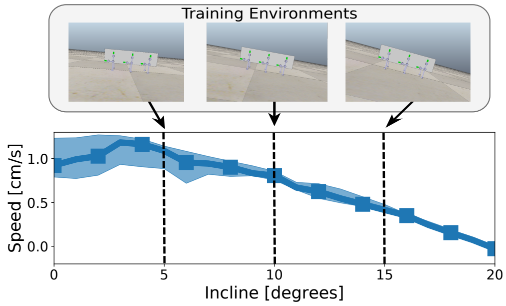
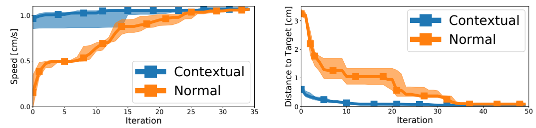
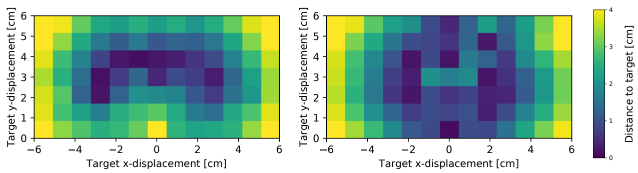


Figure 5: Performance of the contextual policy (median and 65th percentile) for a wide range of inclines.



(a) Inclined surface. (b) Curved trajectory.

Figure 6: Comparison between the optimization performance of a contextual optimizer and a normal optimizer for two different tasks. The contextual optimizer can leverage prior experiments to obtain high-performing gaits in fewer experiments.



(a) cBO (b) Our approach

Figure 7: Comparison of the performances of cBO and our approach for learning motor primitives (using the same data). Darker color indicates better target accuracy.

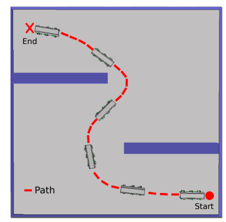


Figure 8: Path constructed using the locomotion primitives learned with our approach.

inclines is time-consuming, so we trained on: 5, 10, and 15 degrees. After optimizing Dual Tripod for these three inclines over 50 iterations, we tested cBO's generalization across a wider context space. In Figure 5 we see good performance in intermediary inclines and smooth interpolation between the training inclines. As shown in Figure 6a, cBO converged on optimal performance much faster than normal BO. This demonstrate cBO's ability to use data accumulated in other contexts to quickly reach optimality in unseen contexts.

Learning to Curve Another useful task that can be framed as contextual optimization is learning motor primitives to walk curves for use in path planning. In addition to the 4 parameters used in our initial experiments, our 2 context parameters here were target displacements along both the x and y axes from point of origin. To train trajectories, we selected five evenly spaced target points in the field of vision. The objective is to reach the desired destination, so our objective function is the distance of the final position to the target position. Over 10 repetitions, the walker could accurately move and turn towards all of the target points within 250 iterations.

Learning Motor Primitives for Path Planning In the previous experiment we learned motor primitives capable of walking curved trajectories. We now demonstrate how our approach presented in Section 4 can be used to significantly improve movement accuracy (compared to cBO using the same data), as well as how such motor primitives can be used in path planning. First, we reused the data from the previous experiment to reformulate the task as a multi-objective optimization as described in Section 4. Then, we used our trained model to randomly sample 10,000 trajectories from the parameter space. Out of these trajectories, we selected the one with smallest expected error. Evaluating the resulting sequence of motor primitives on the real system (i.e., the simulator) demonstrated that the expected trajectory was capable of navigating the maze.

6 CONCLUSIONS

Designing controllers for locomotion can be daunting. In this paper, we demonstrated on a simulated microrobot that this process can be much automated using BO. Our main contributions are two-fold: 1) we introduced a coherent curriculum of increasing challenging tasks, which we use to evaluate our microrobot using existing BO techniques. 2) we presented a new approach that enables walking robots to efficiently learn motor primitives from scratch. By using the data collected from cBO we reformulate the problem into a MOO task, and learn a model that maps any set of parameters to a predicted trajectory. Our experimental results demonstrate using this approach the microrobot can successfully navigate a maze.

References

[1] J. Branke, K. Deb, K. Miettinen, and R. Słowiński. *Multiobjective optimization: interactive and evolutionary approaches*, volume 5252. Springer, 2008. doi: 10.1007/978-3-540-88908-3.

[2] R. Calandra, A. Seyfarth, J. Peters, and M. P. Deisenroth. Bayesian optimization for learning gaits under uncertainty. *Annals of Mathematics and Artificial Intelligence (AMAI)*, 76(1):5–23, 2015. ISSN 1573-7470. doi: 10.1007/s10472-015-9463-9.

[3] R. Campos, V. Matos, and C. Santos. Hexapod locomotion: A nonlinear dynamical systems approach. In *Annual Conference of IEEE Industrial Electronics Society (IECON)*, pages 1546–1551, Nov 2010. doi: 10.1109/IECON.2010.5675454.

[4] G. Capi, M. Yokota, and K. Mitobe. A new humanoid robot gait generation based on multiobjective optimization. In *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 450–454, 2005. doi: 10.1109/AIM.2005.1501032.

[5] S. Chernova and M. Veloso. An evolutionary approach to gait learning for four-legged robots. In *International Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 2562–2567. IEEE, 2004.

[6] D. S. Contreras, D. S. Drew, and K. S. Pister. First steps of a millimeter-scale walking silicon robot. In *International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, pages 910–913. IEEE, 2017.

[7] D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Journal of Global Optimization*, 21(4):345–383, 2001.

[8] J. Knowles. ParEGO: A hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation*, 10(1): 50–66, January 2006.

[9] H. J. Kushner. A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise. *Journal of Basic Engineering*, 86:97, 1964.

[10] D. J. Lizotte, T. Wang, M. Bowling, and D. Schuurmans. Automatic gait optimization with Gaussian process regression. In *International Joint Conference on Artificial Intelligence (IJCAI)*, pages 944–949, 2007.

[11] J. H. Metzen, A. Fabisch, and J. Hansen. Bayesian optimization for contextual policy search. In *IROS Workshop on Machine Learning in Planning and Control of Robot Motion*, 2015.

[12] C. Niehaus, T. Röfer, and T. Laue. Gait optimization on a humanoid robot using particle swarm optimization. In *Proceedings of the Second Workshop on Humanoid Soccer Robots in conjunction with the*, 2007.

[13] M. Oliveira, L. Costa, A. Rocha, C. Santos, and M. Ferreira. Multiobjective optimization of a quadruped robot locomotion using a genetic algorithm. In *Soft Computing in Industrial Applications*, volume 96, pages 427–436. Springer, 2011. ISBN 978-3-642-20504-0. doi: 10.1007/978-3-642-20505-7_38.

[14] M. Oliveira, V. Matos, C. P. Santos, and L. Costa. Multi-objective parameter CPG optimization for gait generation of a biped robot. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 3130–3135, 2013.

[15] V. Pareto. *Manuale di Economia Politica*, volume 13. Societa Editrice, 1906.

[16] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[17] K. Saito, K. Iwata, Y. Ishihara, K. Sugita, M. Takato, and F. Uchikoba. Miniaturized Rotary Actuators Using Shape Memory Alloy for Insect-Type MEMS Microrobot. *Micromachines*, 7 (4), 2016. ISSN 2072-666X. doi: 10.3390/mi7040058.

[18] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.

[19] R. Tedrake, T. W. Zhang, and H. S. Seung. Stochastic policy gradient reinforcement learning on a simple 3d biped. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 2849–2854. IEEE, 2004.

[20] M. Tesch, J. Schneider, and H. Choset. Using response surfaces and expected improvement to optimize snake robot gait parameters. In *International Conference on Intelligent Robots and Systems (IROS)*, pages 1069–1074. IEEE, 2011.

[21] M. Tesch, J. Schneider, and H. Choset. Expensive multiobjective optimization for robotics. In *International Conference on Robotics and Automation (ICRA)*, pages 973 – 980, 2013.

[22] D. Vogtmann, R. S. Pierre, and S. Bergbreiter. A 25 mg magnetically actuated microrobot walking at >5 body lengths/sec. In *IEEE International Conference on Micro Electro Mechanical Systems (MEMS)*, pages 179–182, Jan 2017. doi: 10.1109/MEMSYS.2017.7863370.

[23] T. Wagner, M. Emmerich, A. Deutz, and W. Ponweiser. On expected-improvement criteria for model-based multi-objective optimization. In *Parallel Problem Solving from Nature (PPSN) XI*, pages 718–727. Springer, 2010.

[24] J. Yu, M. Tan, J. Chen, and J. Zhang. A survey on CPG-inspired control models and system implementation. *IEEE Transactions on Neural Networks and Learning Systems*, 25(3):441–456, mar 2014. doi: 10.1109/tnnls.2013.2280596.