Rapid Discovery of Legged Robot Designs via Model Predictive Control

Mihika Kamat[1]

***Abstract*— Through the process of evolution, natural organisms have developed distinct morphologies suited to their environment and mode of survival. With this as inspiration, recent work has studied generating robot designs de novo with evolutionary algorithms. To measure fitness, these works commonly use reinforcement learning (RL) to first train a controller for a given task and robot design. Although this has shown success, evolution with RL is costly, and can be computationally prohibitive when searching over a large population size or scaling the number of generations. In this work, we introduce Evolutionary Predictive Sampling (EPS): an evolutionary algorithm that uses a derivative-free variant of model predictive control (MPC) in the inner loop. Unlike RL, MPC requires no prior training on a fixed morphology, making it uniquely suited to evaluating candidate designs with minimal simulation. This enables our method to use 80x fewer simulation steps per morphology evaluation. On the task of rough terrain locomotion, EPS achieves 63% higher task reward compared to an RL baseline with a fixed compute budget of 50 billion simulation steps. We evaluate EPS on 3 additional tasks: walking on flat ground, reorienting a cube, and transporting an object to a goal, where the method produces consistent gains in population fitness over 50 generations. We believe our work will help accelerate future research in robot co-design by significantly reducing the compute and wall-clock time required to compare a large number of morphologies.**

I. Background

Beginning with Karl Sims's foundational work in robot design discovery in 1994 [3], research in robot design has overwhelmingly used design search algorithms based on morphology-controller co-design, where a morphology and controller is learned simultaneously. Within co-design, popular techniques such as the usage of universal "zero-shot" controllers that may later be fine-tuned [4, 6]; asynchronous design evaluation; and cluster-universal controllers that are used across structurally similar robot designs [1]; have been used to optimize compute and wall-clock for robot design discovery. Even with these changes, co-design remains computationally expensive and unsustainable for rapid design discovery.

Other research such as [5] opts against RL, using a derivative-free implementation of Model Predictive Control: MPPI, for quick, compute-efficient evaluation of robot designs. In this work we adapt this idea, adopting the usage of another derivative-free implementation of Model Predictive Control: Predictive Sampling [2] for design evaluation. Similar to [1], we model robot designs with kinematic trees, rendered and simulated on MuJoCo software, adapting [2]'s existing Predictive Sampling algorithm for efficient design-evaluation.

*This work was not supported by any organization

[1]Mihika Kamat is a student at Los Altos High School, Los Altos, CA 94022, USA mihika.kamat@gmail.com

II. Methodology

A. Evolutionary Predictive Sampling

In this work, we present Evolutionary Predictive Sampling (EPS), an evolutionary algorithm that uses predictive sampling for rapid design evaluation. Predictive Sampling [2] performs periodic replanning during simulation by sampling candidate control sequences from a Gaussian distribution centered on the current controller, projected H seconds into the future, and selecting the highest-performing sequence under the objective. Each replanning step comprises multiple successive sampling layers where the sampling distribution is recentered around the best-performing sequence identified thus far, and additional samples are drawn to iteratively refine the solution. We begin with a population of 128 procedurally generated legged robot designs. Each morphology starts with a head body, and receives a randomly chosen sequence of mutations adapted from [1]. At each generation, every member of the population is evaluated for fitness using 10 trajectories, with actions sampled using the predictive sampling algorithm. The best performing morphology per group of 4 is mutated and used to replace the oldest members of the population. Further details are outlined in Algorithm 1.

Algorithm 1 Evolutionary Predictive Sampling (EPS)

Require: initial population $\mathcal{P}$, objective J

Evolve: (repeat)
1: **Evaluate** each design in population $\mathcal{P}$ for task fitness
2: Randomize groups of size k
3: Mutate best performing design in each group and insert into $\mathcal{P}$
4: Delete $|\mathcal{P}|/k$ oldest members in population

Evaluate: (repeat)
1: **Plan** to update policy $\mathbf{\Pi}(\theta)$
2: Read the current action u from policy $\mathbf{\Pi}(\theta)$, apply it to the controlled system
3: Accumulate J to compute fitness

Plan: (repeat)
1: Measure the current state x
2: Get $N-1$ samples $\theta^{(i)} \sim \mathcal{N}(\theta, \sigma^2)$
3: Including θ, roll out all N samples from x
4: Pick the best one: $\theta \leftarrow \text{argmax}\big(J(\theta^{(i)})\big)$

B. Comparison to RL baseline

We create an RL baseline using for a rough terrain locomotion task for comparison with EPS. The baseline replaces the predictive sampling evaluation step in EPS with a two-step controller-training and design-evaluating method, run for every design. We train designs by using Proximal Policy Optimization (PPO) with 2048 parallel environments and 300 iterations. To reduce wall-clock time, scores for designs already evaluated are cached for future design selection.

C. Additional Tasks

In addition to rough terrain locomotion, we evaluate EPS on 3 more tasks: flat terrain locomotion, manipulating a cube to a goal, and reorienting a cube to a goal orientation.

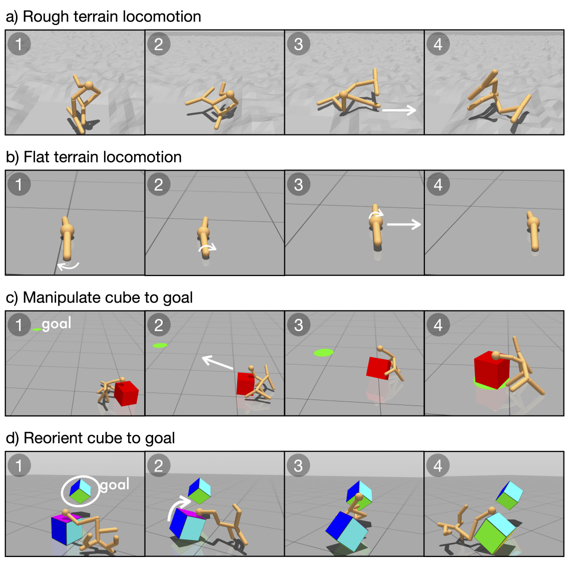

Fig. 1. Learned Morphologies - trajectories for all tasks: a) rough-terrain locomotion, b) flat-terrain locomotion, c) manipulate cube to goal, d) reorient cube to goal

III. Conclusion

Overall, EPS successfully discovers unique designs in the legged robot design space that show high performance in their respective of tasks. In terms of efficiency, it is a suitable alternative to RL-based evolution methods, allowing for accurate discovery, while also achieving 80x more evaluation throughput as measured in designs evaluated per simulation steps.

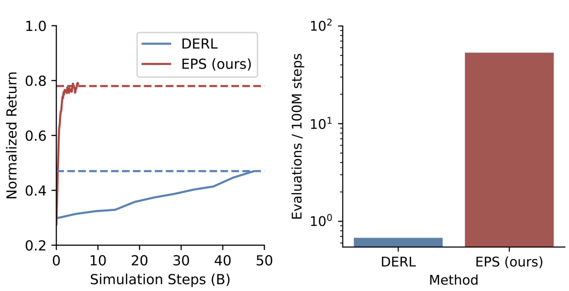

Fig. 2. Comparing EPS to RL baseline

For all four tasks, there is an increase in design fitness across 50 generations. Diversity tends to reduce in terms of skeletal structure, but the population continues to exhibit variations in limb length and other physical properties.

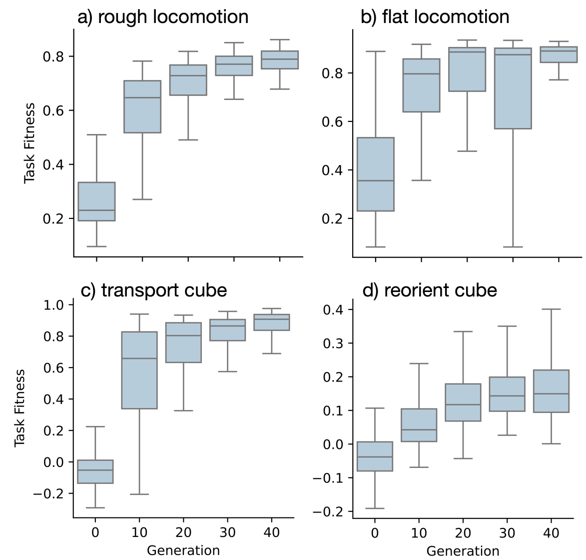

Fig. 3. Distribution of fitness score for morphologies versus evolution generations completed for each task.

IV. Limitations

Future work could focus on addressing the following areas:

- **Overcoming sparse rewards.** Predictive sampling relies on dense rewards, due to a finite planning horizon. This means that tasks that require a longer time-frame or have more binary reward functions, can be harder to evaluate. Future work could explore the use of value functions to overcome these limitations.
- **Hybrid Tasks.** To make EPS useful for real world robot design, morphologies should be evaluated on multiple tasks simultaneously.
- **Early Diversity loss.** This may be a sign of reaching a local optima in the design space. Additional tests for design elimination may be required to prevent the early propagation of designs that are too similar.
  Future research should focus on addressing these limitations so that EPS can be used for a larger set of tasks, especially those needed for the real world.
  - Parallelization - have to rely on cpu-based simulation.
  - short-sightedness

Acknowledgment

Thanks to Adam Rashid and Darren Dressen for mentorship on this project, and Sarthak Kamat for your feedback on this draft. Thanks to National Research Platform (NRP) for cluster access.

References

[1] Gupta, A., Savarese, S., Ganguli, S., et al. (2021). Embodied intelligence via learning and evolution. Nature Communications, 12, 5721.

[2] Howell, T., Gileadi, N., Tunyasuvunakool, S., Zakka, K., Erez, T., & Tassa, Y. (2022). Predictive Sampling: Real-time Behaviour Synthesis with MuJoCo. arXiv:2212.00541.
[3] Sims, K. (1994). Evolving virtual creatures. In Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '94) (pp. 15–22). Association for Computing Machinery.
[4] Strgar, L., & Kriegman, S. (2025). Accelerated co-design of robots through morphological pretraining. arXiv:2502.10862.
[5] Allan Zhao, Jie Xu, Mina Konaković-Luković, Josephine Hughes, Andrew Spielberg, Daniela Rus, and Wojciech Matusik. 2020. RoboGrammar: graph grammar for terrain-optimized robot design. ACM Trans. Graph. 39, 6, Article 188 (December 2020), 16 pages.
[6] A. Gupta, L. Fan, S. Ganguli, and L. Fei-Fei, "MetaMorph: Learning Universal Controllers with Transformers ," in Proc. 10th Int. Conf. Learn. Represent. (ICLR), 2022.
[7] Pathak, D., Lu, C., Darrell, T., Isola, P., & Efros, A. A. (2019). Learning to Control Self-Assembling Morphologies: A Study of Generalization via Modularity. arXiv:1902.05546.
[8] Bhatia, J. S., Jackson, H., Tian, Y., Xu, J., & Matusik, W. (2022). Evolution Gym: A Large-Scale Benchmark for Evolving Soft Robots. arXiv:2201.09863.
[9] Ha, D. (2019). Reinforcement Learning for Improving Agent Design. Artificial Life, 25(4), 352–365.