Efficient Multi-Objective Reinforcement Learning with Expressive Function Approximators

Author: **Adam Štafa**
Supervisor: Petr Novotný

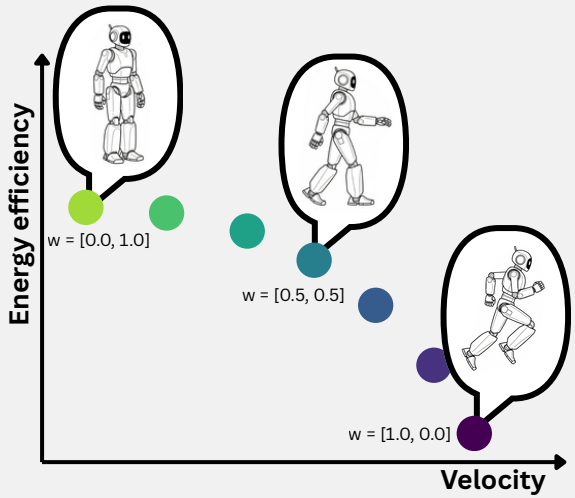
MUNI FI

Reinforcement Learning

Reinforcement learning (RL) is an area of machine learning focused on sequential decision making. The goal is to train a controller that, based on observations, proposes actions to maximize a reward signal. RL is used in e.g. robotics, LLM post-training, and game playing.

Multi-Objective RL

Real-world problems often require optimizing multiple conflicting objectives simultaneously. Multi-objective reinforcement learning (MORL) learns a Pareto front of controllers, each optimizing a different trade-off. The difficulty lies in representing this wide range of different behaviors and learning them efficiently.



Robot locomotion: *Maximize speed while minimizing energy consumption*

Methodology

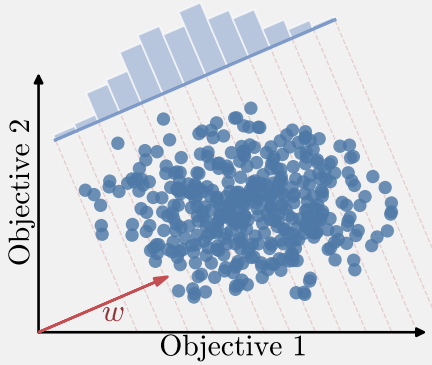
Recent MORL research has mostly focused on algorithmic innovations while using simple feedforward neural networks. We take a different approach, by starting with CAPQL as our base and improving its function approximation, resulting in the Momba algorithm.

1. Neural Architectures

The algorithm fits multiple different behaviors in a single neural network. We thus employ the SimbaV2 architecture to increase the learning capacity of the policy and value networks.

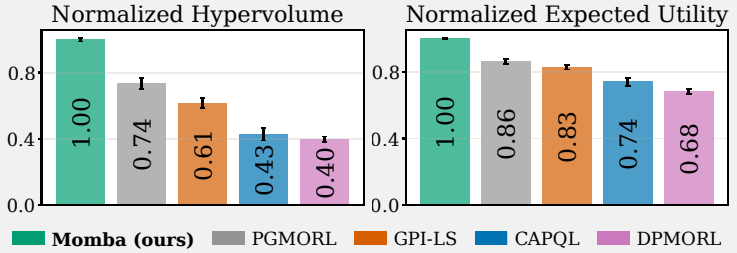
2. Distributional Learning

The agent learns to predict the rewards it will obtain. While MORL algorithms commonly learn the mean reward, we instead adapt distributional value representation to multi-objective RL and stabilize the learning dynamics.

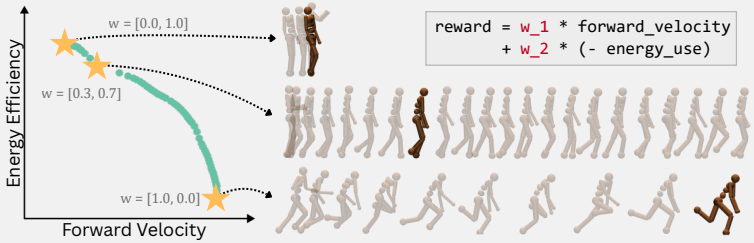


Evaluation and Results

We evaluated Momba on 7 continuous control environments, each with 2 or 3 objectives. The algorithm substantially outperformed other MORL algorithms in Pareto front quality and coverage, as measured by hypervolume and expected utility.



Even in the challenging high-dimensional humanoid environment, Momba recovered the whole Pareto front and discovered useful behaviors, such as standing, walking, and running.



Achievements

- SVOČ 2026: 1st place in the AI category
- Peer-reviewed publication in the Reinforcement Learning Journal, presented at RLC 2026

Adam Štafa, Santeri Heiskanen, Petr Novotný, & Joni Pajarinen (2026). Momba: Network Modernization Improves Multi-Objective Reinforcement Learning. Reinforcement Learning Journal, 7.