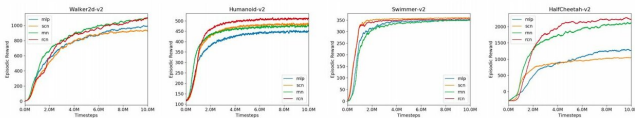


A new approach for modeling central pattern generators (CPGs) in reinforcement learning

23 January 2019, by Ingrid Fadelli



Plots comparing the baseline models (MLP, SCN, RNN, RCN) for the 4 MuJoCo environments presented in the paper (Humanoid-v2, HalfCheetah-v2, Walker2d-v2, Swimmer-v2). Credit: Liu et al.

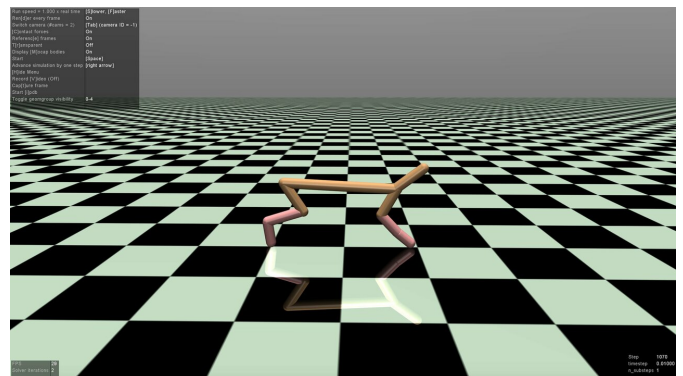
Central pattern generators (CPGs) are biological neural circuits that can produce coordinated rhythmic outputs without requiring rhythmic inputs. CPGs are responsible for most rhythmic motions observed in living organisms, such as walking, breathing or swimming.

Tools for effectively modeling rhythmic outputs when given arrhythmic inputs could have important applications in a variety of fields, including neuroscience, robotics and medicine. In [reinforcement learning](#), most existing networks used to model locomotive tasks, such as multilayer perceptron (MLP) baseline models, fail to generate rhythmic outputs in the absence of rhythmic inputs.

Recent studies have proposed the use of architectures that can split a network's policy into linear and nonlinear components, such as structured control nets (SCNs), which were found to outperform MLPs in a variety of environments. An SCN comprises a linear model for [local control](#) and a nonlinear module for global control, the outputs of which are combined to produce the policy action. Building on prior work with recurrent neural networks (RNNs) and SCNs, a team of researchers at Stanford University has recently

devised a new approach to model CPGs in reinforcement learning.

"CPGs are biological neural circuits capable of producing rhythmic outputs in the absence of rhythmic input," Ademi Adeniji, one of the researchers who carried out the study, told Tech Xplore. "Existing approaches for modeling CPGs in reinforcement learning include the multilayer perceptron (MLP), a simple, fully-connected neural network, and the structured control net (SCN), which has separate modules for local and global control. Our research objective was to improve upon these baselines by allowing the model to capture previous observations, making it less prone to error from input noise."



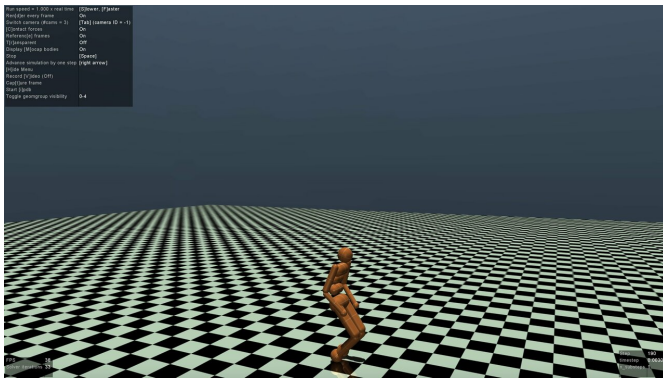
Screenshot of the HalfCheetah environment. Credit: Liu et al.

The recurrent control net (RCN) developed by Adeniji and his colleagues adopts the architecture of a SCN, but uses a vanilla RNN for global control. This allows the model to acquire local, global and time-dependent control.

"Like SCN, our RCN splits the information flow into linear and nonlinear modules," Nathaniel Lee, one of the researchers who carried out the study, told TechXplore. "Intuitively, the linear module, effectively a linear transformation, learns local interactions, whereas the nonlinear module learns global interactions."

SCN approaches use an MLP as their nonlinear module, while the RCN devised by the researchers replaces this module with a RNN. As a result, their model acquires a 'memory' of past observations, encoded by the RNN's hidden state, which it then uses to generate future actions.

The researchers evaluated their approach on the OpenAI Gym platform, a physics environment for reinforcement learning, as well as on multi-joint dynamics with contact (MuJoCo) tasks. Their RCN either matched or outperformed other baseline MLPs and SCNs in all tested environments, effectively learning local and global control while acquiring patterns from prior sequences.



Screenshot of the Humanoid environment. Credit: Liu et al.

"CPGs are responsible for a vast number of rhythmic biological patterns," Jason Zhao, another researcher involved in the study, said. "The ability to model CPG behavior can be successfully applied to fields such as medicine and robotics. We also hope that our research will highlight the efficacy of local/global control as well as recurrent

architectures for modeling central pattern generation in reinforcement learning."

The findings gathered by the researchers confirm the potential of SCN-like structures to model CPGs for reinforcement learning. Their study also suggests that RNNs are particularly effective for modeling locomotive tasks and that separating linear and nonlinear control modules can significantly improve a model's performance.

"So far, we only trained our model using evolutionary strategies (ES), an off-gradient optimizer," said Vincent Liu, one of the researchers involved in the study. "In the future, we plan to explore its performance when training it with proximal policy optimization (PPO), an on-gradient optimizer. Additionally, advances in natural language processing have shown that convolutional neural networks are effective substitutes to recurrent neural networks, both in performance and computation. We could hence consider experimenting with a time-delay neural network architecture, which applies 1-D convolution along the time axis of past observations."

More information: Recurrent control nets for deep reinforcement learning. arXiv:1901.01994 [cs.LG]. arxiv.org/abs/1901.01994

© 2019 Science X Network

APA citation: A new approach for modeling central pattern generators (CPGs) in reinforcement learning (2019, January 23) retrieved 20 June 2021 from <https://techxplore.com/news/2019-01-approach-central-pattern-cpgs.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.