Using Training Modules To Train an Agent

1st Tofara Moyo
Bulawayo , Zimbabwe
tofaramoyo@gmail.com, Mazusa AI

***Abstract*—We propose a novel framework for training humanoid robots to exhibit human-like behavior by leveraging musical consonance as a guiding principle. After indexing the neurons in a spiking neural network with the names of keys in a musical keyboard it is trained to produce consonant activations in response to human-generated data, while simultaneously learning to distinguish between human-like and robot-like behavior by producing dissonant activations in response to robot generated data.Then, through reinforcement learning, a humanoid robot is trained to mimic human behavior by using consonance in the network's activations as a reward while the network is shown the robots generated data. Our approach enables the development of modular, task-specific skills ,one per spiking network, and demonstrates the potential for scalable and flexible behavioral learning in humanoid robots.**

I. Introduction

This paper presents a novel methodology for developing modular learning components, herein referred to as 'learning modules,' designed for integration into autonomous agents. These modules enable agents to rapidly acquire specific skills or tasks upon loading, leveraging the principles of modularity to facilitate flexibility, scalability, and efficiency. By compartmentalizing learning processes into discrete, reusable modules, our approach offers numerous benefits, including enhanced adaptability, reduced training times, and improved maintainability.

Modularity in systems refers to the design principle of breaking down a complex system into smaller, independent, and interchangeable modules or components. Each module performs a specific function, and these modules can be easily combined, rearranged, or replaced without affecting the entire system.

Benefits of modularity include:

Flexibility: Modules can be easily added, removed, or modified as needed.

Scalability: Systems can be scaled up or down by adding or removing modules.

Maintainability: Faulty modules can be quickly replaced, reducing downtime.

Reusability: Modules can be reused across different systems or projects.

Easier debugging: Issues can be isolated to specific modules. Improved collaboration: Different teams can work on separate modules simultaneously.

This design approach is commonly applied in software development, electronics, and other fields where complex systems are built from smaller components.Here we use it in a reinforcement learning setting to train an agent to be able to perform multiple tasks.

We propose a framework centered on training a neural network architecture. Specifically, we hypothesize that a spiking neural network (SNN) would be an optimal choice for this application, owing to its discrete, event-driven activation mechanism

A spiking neural network (SNN) is a type of artificial neural network inspired by the biological neural networks in the brain. Unlike traditional neural networks that process information using continuous signals, SNNs mimic the brain's neural activity by transmitting information through discrete events, or "spikes," in time.

In SNNs:

Neurons communicate through spikes, which are sudden increases in membrane potential.

The timing and frequency of spikes convey information. Synaptic weights and plasticity rules govern learning and adaptation.

The spiking neural network will undergo task-specific training, wherein the training dataset is carefully curated to comprise tasks that align with the intended functionality of the module. This targeted training approach aims to optimize the network's performance and adaptability within its designated domain.

To train the spiking neural network, we draw inspiration from the work of Moyo et al., who propose a novel regularization technique utilizing music grammar-based activations.

Specifically, their approach leverages a plugin commonly employed to quantify perceptual consonance, applying it to measure the consonance of network layers.

Specifically, the neurons in each layer were indexed according to a musical keyboard paradigm, where each neuron was associated with a specific key. At each time step, the subset of neurons that fired represented a particular chord, allowing the plugin to quantify the consonance level of the resulting neural activity pattern.

The consonance level quantified by the plugin serves as the reward signal for the reward-modulated spiking neural network.

This protocol led to the emergence of structured, music-like chord patterns within the neural network. To further shape the network's behavior, we introduced an additional reward mechanism that incentivized smooth voice leading, specifically favoring chord transitions characterized by intervals of perfect fourths or fifths between consecutive time steps.

We propose a framework for capturing human behavior and activity through multimodal sensing, utilizing wearable cameras and microphones to record a human subject's actions and interactions. The ultimate objective is to transfer the learned patterns and behaviors to a humanoid robot, enabling it to mimic the human's actions and interactions with high fidelity.

We adopt a training paradigm inspired by generative adversarial networks (GANs), wherein a spiking neural network is trained on human-generated data to produce consonant activations adhering to musically meaningful patterns (specifically, motion in fourths or fifths). Concurrently, we record the behavior of an untrained humanoid robot via wearable cameras and microphones, and train the same network to generate dissonant, musically unmeaningful activations. This adversarial training framework enables the network to learn a representation that distinguishes between human-like, consonant behavior and robot-like, dissonant behavior.

Upon completion of the training protocol, the module network would be deployed to guide the behavior of the humanoid robot. To visualize the network's activity, we would generate a real-time simulation of the spiking neurons, utilizing a color-coded representation to denote different notes/neurons. This visual representation would be projected onto a designated region of the robot's visual input, enabling the integration of the network's musical patterns with the robot's sensory perception.

The agent's training will be completed through reinforcement learning, utilizing the consonance level of the spiking neural network's activations as the reward signal. Simultaneously, the spiking network will receive real-time input from the robot's sensors, allowing the agent to learn a policy that makes it more human like.

This framework enables the robot to learn human-like behavior through an implicit objective: maximizing consonance in the module network's activations. As a result, the robot will adapt its actions to increasingly mimic the patterns and behaviors exhibited by the human in the training data, thereby inducing harmonious activations in the network.

This approach offers the flexibility to develop multiple modules, which can be integrated and trained on the humanoid robot either concurrently or sequentially, allowing for a modular and scalable framework for behavioral learning.

Potential applications of this framework could include the development of specialized modules tailored to specific skills, such as fine motor control, conversational dialogue, or task-specific expertise (e.g., construction or manufacturing). This modular approach would enable the humanoid robot to acquire a diverse range of abilities, each optimized for a particular domain or function.

II. Conclusions

In conclusion, our study demonstrates the feasibility of using musical consonance as a guiding principle for training humanoid robots to exhibit human-like behavior. By leveraging the structural properties of music, we have developed a novel framework that enables robots to learn complex patterns and behaviors from human data. The modular and scalable nature of our approach offers promising avenues for future research, including the development of task-specific skills and the integration of multiple modules for more comprehensive behavioral learning. Our findings suggest that the intersection of music and robotics can provide valuable insights into the design of more sophisticated and human-like artificial intelligence systems.

References

[1] Generative Adversarial Networks Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio https://arxiv.org/abs/1406.2661

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, pp. 436–444, 2015.

[3] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in International conference on machine learning. PMLR, 2015, pp. 1889–1897.

[4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," arXiv preprint arXiv:1707.06347, 2017.

[5] Structuring Concept Space with the Musical Circle of Fifths by Utilizing Music Grammar Based Activations-Tofara Moyo https://arxiv.org/abs/2403.00790