RESEARCH ARTICLE OPEN ACCESS

EXPLAINABLE NEURO-EVOLUTIONARY REINFORCEMENT LEARNING FOR TRUSTWORTHY AUTONOMOUS DECISION SYSTEMS

R. Abiramipriya*, N. Dhanush, M. Mohamed Ajmal**, A. Alwin Joel ****
* Assistant Professor -Department of Computer Science and Engineering,
** UG Students- Department of Computer Science and Engineering,
A.V.C. College of Engineering,
Mannampandal, Mayiladuthurai, India

Abstract
Deep reinforcement learning has exhibited good results in autonomous control and sequential decision-making. Nevertheless, its application to safety-critical settings is still limited by the fact that learned policies have little interpretability. In actual autonomous systems, an agent is not only required to make effective decisions, but also offer clear and verifiable reason as to why it has done so. In order to overcome this issue, this paper suggests an Explainable Neuro-Evolutionary Reinforcement Learning (XNE-RL) model which integrates the policy evolution process, multi-objective optimization, and interpretability in a single learning system. The suggested model develops a population of neural policies via genetic manipulation of selection, cross over and mutation and at the same time maximizes the expected return, policy complexity as well as the quality of the explanation. Besides that, there is also a real-time explainability module that identifies the salient input variables, tracing intermediate activations, and deriving human-readable decision rules when executing policy. It is tested on autonomous navigation and multi-agent coordination objectives and compared to Deep Q-Network (DQN) and Proximal Policy Optimization (PPO) baselines. The suggested solution should bring about competitive decision performance with less complex model and high interpretability hence supporting the creation of credible autonomous decision systems.

Keywords: Explainable artificial intelligence; reinforcement learning; neuro-evolution; multi-objective optimization; autonomous systems; interpretable policy learning

1. Introduction

Reinforcement learning has emerged as a significant paradigm in the autonomous systems, robotics and intelligent control of solving sequential decision making problems. The latest developments in the deep reinforcement learning have however advanced the ability of agents to learn complex policies directly using high dimensional sensory inputs. Although such accomplishments have been realized, the lack of insight into the workings of deep neural policies has been a significant constraint in areas that require safety, accountability, and trust.

The autonomous decision systems applied to transportation, surveillance, healthcare, and industrial automation have to justify their actions in a way that will be understandable by humans. The traditional methods of reinforcement learning primarily maximize the performance goals,

e.g. cumulative reward, without directly addressing the level of explainability of a policy behavior. This could make it hard to validate, audit, or trust outputs of such systems by the decision-makers.

In order to address this concern, this paper presents an Explainable Neuro Evolutionary Reinforcement Learning (XNE-RL) framework. The point is to make explainability a part of the policy optimization process instead of considering it a post hoc analysis step. The presented framework is a neuro-evolutionary search with multi-objective optimization aimed at ensuring an overall improvement of three aspects: policy effectiveness, model simplicity, and interpretability.

The principal contributions of this work are the following:

An XNE-RL framework of a reliable autonomous decision-making.

Multi-objective optimization approach to ensure a tradeoff between reward and complexity with the quality of the explanation.

An explainability component that can be used in real-time to extract key input variables and produce interpretable decision rules.

A multi-agent coordination environment with autonomous navigation evaluation strategy.

Comparison to the traditional reinforcement learning baselines such as DQN and PPO.

The rest of the paper is structured in the following way. Related work is reviewed in section 2. Section 3 is the statement of the problem. The proposed methodology is offered in Section 4. The description of the experimental design is presented in Section 5. The projected outcomes and analysis are discussed in Section 6. Section 7 concludes the paper.

2. Related Work

Deep reinforcement learning has recorded a significant success in the field of control, planning, and game-playing. The algorithms like the DQN and PPO have demonstrated a high ability to learn in diverse benchmark settings. Nevertheless, such approaches tend to create neural policies which are in interpretable nature. Explainable AI has grown to become a significant field of research that aims to enhance transparency in the machine learning systems. The current explainability approaches usually use feature attribution, saliency visualization, surrogate models, or rule extraction. Although they can be helpful, these approaches are usually used after the training and in practice are not relevant to the actual decision mechanism of the policy that can take place during its functioning.

Another potentially successful direction in research is neuro-evolution where neural networks are optimized by evolutionary methods, not just using the gradient approach. Such problems as the investigation of policy structures, the management of sparse rewards and the optimization of multiple objectives are best addressed using evolutionary methods. Nevertheless, the majority of neuro evolution approaches focus on the task performance and explicitly do not consider the explainability.

The suggested XNE-RL framework is based on the three areas and integrates reinforcement learning, neuro-evolution, and explainable AI into a single framework. This integration facilitates learning of not only effective policies but also easier to comprehend and prove.

3. Problem Formulation

Consider an environment modeled as a Markov Decision Process defined by the tuple

$$\mathcal{M} = (\square\square\square\square\square)$$

where ◻◻denotes the state space, ◻◻denotes the action space, ◻◻represents the transition dynamics, ◻◻is the reward function, and ◻◻∈ [0,1] is the discount factor. The objective of reinforcement learning is to identify a policy ◻◻(◻◻| ◻◻) that maximizes the expected cumulative discounted reward:

$$◻◻(◻◻) = ◻◻[\sum_{◻◻=0}^{◻◻} ◻◻^{◻◻} ◻◻◻◻]$$

In the proposed work, policy learning is extended beyond reward maximization by introducing two additional objectives:

1. **Policy complexity minimization**
2. **Explanation quality maximization**

The optimization problem is therefore formulated as a multi-objective problem: $\max ◻◻◻◻(◻◻) = [◻◻(◻◻), −◻◻(◻◻), ◻◻(◻◻)]$

where:

- ◻◻(◻◻) is the expected return,
- ◻◻(◻◻) is the policy complexity,
- ◻◻(◻◻) is the explainability score.

This formulation encourages the learning of policies that achieve high performance, maintain compact structures, and provide interpretable reasoning.

4. Proposed Methodology

4.1 Overview of XNE-RL Framework

The XNE-RL framework consists of three main components:

1. **Neural policy population**
2. **Evolutionary optimization engine**
3. **Real-time explainability module**

A population of neural network policies is initialized and iteratively evolved over multiple generations. Each policy interacts with the environment and is evaluated

Let the population at generation ◻◻ be represented as

$$◻◻^{(◻◻)} = \{◻◻_{1}^{(◻◻)}, ◻◻_{2}^{(◻◻)}, \dots, ◻◻_{◻◻◻◻}^{(◻◻)}\}$$

where ◻◻is the population size. Each policy is encoded by its neural architecture and associated parameters.

4.3 Fitness Evaluation

The fitness of each policy is computed using a composite fitness measure based on reward, complexity, and explainability.

4.2 Population Initialization

using a weighted objective function:

$$◻◻_{◻◻◻} = ◻◻◻◻(◻◻_{◻◻◻}) − ◻◻◻◻(◻◻_{◻◻◻}) + ◻◻◻◻(◻◻_{◻◻◻})$$

where ◻◻, ◻◻ and ◻◻are scalar coefficients controlling the importance of task performance, compactness, and interpretability.

4.4 Evolutionary Operators

The evolution process includes the following operations:

- **Selection:** High-fitness policies are selected for reproduction.
- **Crossover:** Parameter or structural information is exchanged between selected parent policies.
- **Mutation:** Random modifications are introduced to maintain diversity and improve exploration.

These operators generate a new policy population with potentially improved fitness.

4.5 Explainability Module

The explainability component operates during policy execution and includes:

- **Feature importance estimation:** Determines which state variables most strongly influence the chosen action.
- **Activation tracing:** Tracks hidden

layer responses to understand internal policy behavior.

- **Rule extraction:** Converts policy decisions into approximate human readable rules.

The explanation quality score is defined as

$$EQ(\pi) = w_1 \cdot FI + w_2 \cdot AT + w_3 \cdot RR$$

where:

- FI is the feature importance interpretability score,
- AT is the activation transparency score,
- RR is the rule readability score,
- w_1, w_2, w_3 are weighting factors.

4.6 Policy Complexity Measure Policy complexity may be expressed as a function of architectural size:

$$C(\pi) = P + L \cdot S$$

where P is the number of trainable parameters, L is the number of layers, and S is a scaling constant.

5. Algorithm

Algorithm 1: Explainable Neuro Evolutionary Reinforcement Learning

Input: Environment $\mathcal{M}$, population size N, generations G

Output: Optimized explainable policy π^*

1. Initialize a population of neural policies $\pi^{(0)}$
2. For each generation $g = 1$ to G
   i. Evaluate each policy in the environment
   ii. Compute reward $R(\pi)$
   iii. Compute complexity $C(\pi)$
   iv. Compute explainability score $EQ(\pi)$
   v. Calculate overall fitness $F(\pi)$
   vi. Select high-fitness policies
   vii. Apply crossover and mutation to generate offspring
   viii. Form the next generation population
3. Return the best policy

Experimental Design

6.1 Evaluation Environments

The suggested framework is tested on two example tasks:

6.1.1 Autonomous Navigation

An agent is in a dynamic setting with hindrances and target sites. The aim is to get to the target without colliding or as cheaply as possible.

6.1.2 Multi-Agent Coordination

Several agents collaborate to accomplish a common goal including formation control, distributed assignment of tasks or collision-free movement.

6.2 Baseline Methods

The suggested XNE-RL model is compared to the following base-line algorithms:

Deep Q-Network (DQN)

Proximal Policy Optimization (PPO)

6.3 Evaluation Metrics

All methods are judged by the following:

Average cumulative reward

Success rate

Policy complexity

Explanation quality

Convergence speed

Computational cost

6.4 Implementation Settings

The implementation settings that can be used in the final manuscript are as follows:

Population size

Number of generations

Mutation rate

Crossover probability

Learning episodes

The architecture of a neural network.

Hardware and software description.

7. Results and Discussion

It is assumed that the suggested XNE-RL framework will show better or competitive results in cumulative reward and the success rate of tasks compared to DQN and PPO. Meanwhile, its multi-objective design is aimed at generating smaller and easier-to-understand policies.

The explainability module can also be used to validate and debug in tasks of autonomous navigation by showing which sensory inputs have the greatest impact on movement decisions. Under multi-agent coordination environment, the structure can reveal how policy interaction can result in cooperative behavior.

The proposed framework will have the following benefits compared to the usual deep reinforcement learning methods:

Increased openness in course of action.

Reduced policy complexity

Increased safety-critical application.

Improved the level of trust and auditability.

Another possible weakness is the higher amount of calculating power that evolutionary search and explanation generation require. The trade-off can however be explained by the fact that it can be found justifiable in situations where interpretability is a crucial factor. **8. Conclusion**

In this paper, the Neuro-Evolutionary Reinforcement Learning framework that is explained to provide credible autonomous decision-making was introduced. The approach proposed incorporates the policy development, multi-objective optimization, and real-time explainability as one architecture. The framework aims to overcome a significant weakness of traditional deep reinforcement learning in safety-sensitive tasks by jointly maximizing the performance of tasks, structural simplicity, and quality of explanations.

The model also aims at facilitating an effective autonomous action along with transparent decision making, and hence it can be applied in areas that demand accountability and trust. Future efforts can be aimed at supporting the framework with more realistic real world settings, enhancing fidelity to explanations, and decreasing computation time.

10. References

1. Sutton RS, Barto AG. Reinforcement Learning: An Introduction. MIT Press; 2018.
2. Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. Nature. 2015;518:529-533.
3. Schulman J, Wolski F, Dhariwal P, et al. Proximal Policy Optimization Algorithms. arXiv preprint arXiv:1707.06347; 2017.