



RADemics

# Reinforcement Learning Enhanced with Genetic Programming for Real-Time Decision Optimization

B. Angalaparameswari, Senthil Kumar  
Palaniappan, B. Ramesh

RAJALAKSHMI INSTITUTE OF TECHNOLOGY,  
VELALAR COLLEGE OF ENGINEERING AND  
TECHNOLOGY, ADITYA INSTITUTE OF TECHNOLOGY  
AND MANAGEMENT

# Reinforcement Learning Enhanced with Genetic Programming for Real-Time Decision Optimization

<sup>1</sup>B. Angalaparameswari, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Kuthambakkam, Chennai-14, Mail ID: [angalaparameswari.b@ritchennai.edu.in](mailto:angalaparameswari.b@ritchennai.edu.in)

<sup>2</sup>Senthil Kumar Palaniappan, AP-ECE, Velalar College of Engineering and Technology, Erode-12, Mail id: [psenthilec@gmail.com](mailto:psenthilec@gmail.com)

<sup>3</sup>B. Ramesh, Asst. Professor – CSE, Aditya Institute of Technology and Management, Tekkali, Mail id: ramesh. [bandaru15@gmail.com](mailto:bandaru15@gmail.com)

## Abstract

The integration of Reinforcement Learning (RL) with Genetic Programming (GP) presents a robust hybrid framework for real-time decision optimization in dynamic and uncertain environments. Traditional RL methods, while effective in learning optimal policies through interaction, often suffer from limited interpretability, convergence issues, and difficulty adapting to rapidly changing scenarios. Genetic Programming, by contrast, evolves symbolic and human-readable policies, offering transparency and structural flexibility. This chapter investigates the evolution of interpretable policies through GP-driven mechanisms within RL architectures, emphasizing symbolic representation, grammar-guided synthesis, and structural compactness. A detailed exploration is provided on balancing policy performance with interpretability in real-time systems, supported by evaluation frameworks that consider both functional efficacy and cognitive comprehensibility. The role of visualization and explanation tools is examined, particularly in human-in-the-loop systems, to enhance policy transparency, traceability, and auditability. The proposed approach aligns with the growing demand for explainable AI in mission-critical applications, providing a pathway toward intelligent systems that are not only autonomous and adaptive but also accountable and trustworthy.

## Keywords:

Reinforcement Learning, Genetic Programming, Policy Interpretability, Real-Time Decision Optimization, Explainable AI, Symbolic Representation

## Introduction

The rapid advancement of artificial intelligence has led to the increasing reliance on autonomous agents for decision-making in complex, high-speed environments [1]. These agents are often deployed in real-time systems where conditions change rapidly and decisions must be made instantaneously [2]. Reinforcement Learning (RL) has emerged as a powerful paradigm for developing such agents by enabling them to learn optimal policies through continuous interaction

with dynamic environments [3]. RL algorithms use reward signals to evaluate and improve decision-making strategies over time, adapting behavior in response to changing conditions [4]. Its adaptability and theoretical robustness, traditional RL methods face several limitations when applied in real-time systems. Challenges such as high sample complexity, delayed convergence, instability during learning, and lack of interpretability can hinder their performance, particularly in safety-critical and human-in-the-loop applications [5].

To overcome these limitations, researchers have increasingly turned to hybrid approaches that integrate evolutionary computation techniques with reinforcement learning [6]. Among these, Genetic Programming (GP) stands out for its capacity to evolve symbolic, human-understandable representations of policies [7]. Unlike subsymbolic models, such as deep neural networks, GP constructs policies in the form of structured programs, typically represented as trees or logical expressions [8]. These symbolic structures provide inherent transparency and allow decision processes to be understood, audited, and modified with greater ease [9]. GP facilitates global search over the policy space through biologically inspired operators such as crossover and mutation, which helps avoid local optima—a common issue in gradient-based RL algorithms. When combined effectively, RL and GP offer a synergistic framework that balances exploration and exploitation while maintaining model interpretability [10].

This hybridization leverages the strengths of both paradigms: RL contributes an adaptive learning mechanism rooted in real-time interaction, while GP enhances the structure and readability of learned policies [11]. The result is a learning system capable of not only achieving high performance but also producing policies that are traceable and explainable [12]. This is particularly important in real-time decision systems where transparency is as critical as accuracy [13]. For instance, in autonomous driving or industrial process control, system stakeholders must understand and validate the decision-making logic to ensure safety, compliance, and reliability [14]. The symbolic nature of GP-evolved policies addresses this requirement by enabling a direct inspection of the action-selection process. Furthermore, it supports the incorporation of domain knowledge through grammar constraints, which can guide policy evolution within safe and meaningful boundaries [15].