# Multi-objective PIDA controller for the Interpretable Control Competition GECCO 2025


Gilberto Reynoso-Meza
Pontificia Universidade Católica do Paraná, PPGEPS
Curitiba, Paraná, Brazil
g.reynosomeza@pucpr.br
Shinshu University
Nagano, Nagano, Japan

Christian Santiago Feican Campoverde
Pontificia Universidade Católica do Paraná, PPGEPS
Curitiba, Paraná, Brazil
santiago.feican@gmail.com

Nicolas Hupfeld
Pontificia Universidade Católica do Paraná, PPGEPS
Curitiba, Paraná, Brazil
nicolas-hupfeld@hotmail.com

Hernan Aguirre
Shinshu University
Nagano, Nagano, Japan
ahernan@shinshu-u.ac.jp



## Abstract

This paper presents a framework for interpretable learning and control using a PIDA (Proportional-Integral-Derivative-Accelerative) controller in the Atari Pong environment, as part of the GECCO 2025 Interpretable Control Competition. The controller is designed to operate under a discrete action space and uses visual information from the game's rendered frames to identify relevant game objects. We propose a multi-objective optimization of the PIDA parameters to balance three key goals: maximizing match duration, maximizing goals scored, and minimizing goals conceded. The optimization process is performed externally using multi-objective optimization in MATLAB, interfaced via CSV-based communication. The resulting Pareto front enables the exploration of diverse behavioral trade-offs.



## CCS Concepts

• **Operations research → Decision analysis; Multi-criterion optimization and decision-making**.

## Keywords

PIDA controller, multi-objective optimization, interpretable control and learning

**ACM Reference Format:**
Gilberto Reynoso-Meza, Christian Santiago Feican Campoverde, Nicolas Hupfeld, and Hernan Aguirre. 2025. Multi-objective PIDA controller for the Interpretable Control Competition GECCO 2025 . In *Proceedings of GECCO '25*. ACM, New York, NY, USA, 3 pages. https://doi.org/XXXXXXX.XXXXXXX




*GECCO '25, Málaga, SP*



ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX


## 1 Introduction

The intersection of control theory and reinforcement learning has generated increasing interest in interpretable approaches to policy design. Traditional controllers, such as PID-based schemes, offer an intuitive structure and are widely adopted in industrial applications due to their robustness and transparency [1].

This work introduces a PIDA (Proportional-Integral-Derivative-Accelerative) controller for the Atari Pong-v4 environment, designed to operate with only visual input in the form of grayscale frames. We explicitly structure the control loop with interpretable actions (up, down, stay) using the gains to process the error as interpretable inputs (error NOW, PAST error, FUTURE error) by extracting relevant object positions (ball and paddle) through image processing, to approximate y-position of the ball and y-position on the center of mass of the paddle.

The key contribution of this work is the formulation of a multi-objective optimization problem for controller tuning, using evolutionary algorithms in MATLAB to find trade-offs between survival time, offensive performance, and defensive stability to drive the learning process of the gain parameters of the PIDA controller.

## 2 Tools and Methods

The PIDA controller extends the classical PID architecture by incorporating an additional term proportional to the second derivative (acceleration) of the error signal. This structure is particularly useful in systems where higher-order dynamics play a significant role, allowing for faster and more anticipative corrective actions.

In the time domain, the control signal $u(t)$ generated by a PIDA controller is defined as:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau)d\tau + K_d \frac{de(t)}{dt} + K_a \frac{d^2e(t)}{dt^2} \qquad (1)$$

where:

- $e(t) = r(t) - y(t)$ is the control error, with $r(t)$ as the reference signal and $y(t)$ as the measured output,
- $K_p$, $K_i$, $K_d$, and $K_a$ are the proportional, integral, derivative, and accelerative gains, respectively.

To implement closed-loop control, we process the raw RGB observation frames provided by the `Pong-v4` environment to estimate the vertical position of both the agent's paddle and the ball. These positions are extracted using image segmentation techniques in the grayscale domain, applying adaptive thresholding and contour detection. The estimated paddle and ball centroids are then used to compute the control error $e(t)$ as their vertical distance. This error feeds the discrete-time PIDA controller, and it is adapted to operate under a discrete action space, with control decisions mapped to predefined actions (up, down, stay).

The controller gains $\boldsymbol{x} = (K_p, K_i, K_d, K_a)$ are tuned using a multi-objective optimization strategy that considers the performance of each episode in terms of survival time, goals scored, and goals conceded.

$$\min_{\boldsymbol{x}} \boldsymbol{J}(\boldsymbol{x}) = [-J_{frames}(\boldsymbol{x}), -J_{G.scored}(\boldsymbol{x}), J_{G.conceded}(\boldsymbol{x})] \quad (2)$$

subject to:

$$0 \leq x_i \leq 2, i = [1, 2, 3, 4] \quad (3)$$

The optimization process is carried out using the *spMODEx* algorithm[1]. *spMODEx* belongs to the family of spMODE algorithms [3], which have been widely applied and documented in controller tuning tasks in control systems [2]. Finally, principal component analysis will be used to depict in a low dimensional basis the four dimensional landscape of the decision space.

## 3 Results and Discussions

Figure 1 displays the final Pareto front obtained after 500 iterations of the optimization process (population size of 20 vectors), in which the controller gains were tuned through gameplay. Given that both goals scored and goals conceded are discrete variables, the optimization process is susceptible to stagnation. To mitigate this issue and enrich the trade-off space, the number of game frames was incorporated as an additional objective. The inclusion of the frame count serves two critical functions: it prevents premature convergence due to discreteness and guides the optimization process through meaningful trade-offs at different stages of controller development.

Figure 2 depicts the decision space projected onto its first two principal components, derived from the original four decision variables. The plot illustrates the evolution of the offspring proposed in each generation of the algorithm, with their corresponding cost function evaluations. A trajectory line highlights the path followed by the controller $\boldsymbol{x} = [1.8048, 0.0000, 0.7730, 1.1277]$ that maximizes the net score, defined as the sum of goals scored minus goals conceded. This visualization reveals how the search process explores different regions of the decision space and how the selected solution shift over time.

Figure 3a depicts the median value of offspring generated throughout the evolutionary process within the decision variable space, specifically focusing on the four controller gains. Notably, the integral component of the controller converges early in the optimization, which aligns with the characteristics of the process being controlled. As a result, the final controller configuration can be characterized as a PDA-type controller. The green line indicates the generation at which the first winning strategy in the match was discovered, marking a significant milestone in the optimization trajectory. Figure 3b displays the set of parent solutions. These individuals represent the core of the evolving Pareto front and embody the most consistently competitive trade-offs across generations.

[1]Available at https://www.mathworks.com/matlabcentral/fileexchange/65145.

## 4 Conclusions

In this work, a viable controller was proposed for the Interpretable Control Challenge. The initial approach focused on tuning the controller gains through a multi-objective optimization framework. As future work, we aim to explore additional parameters that could be optimized in this game environment, beyond the controller gains. Furthermore, we plan to investigate alternative strategies, including the use of Multi-Objective Reinforcement Learning (MORL), to assess their suitability and effectiveness in handling complex, dynamic, and interpretable control scenarios.

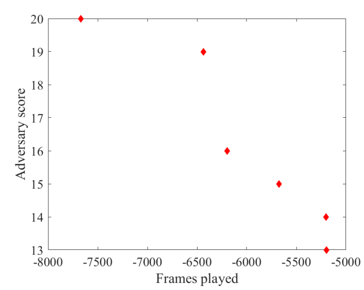



**Figure 1: Pareto front approximation.**

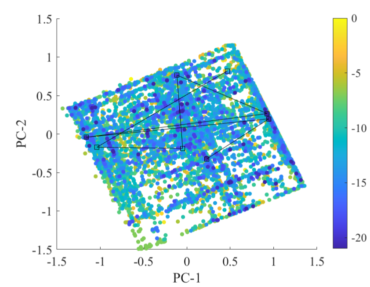
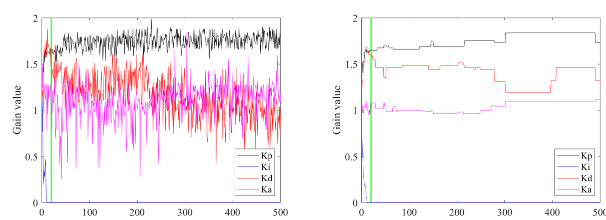



**Figure 2: Landscape representation in 2 principal components.**




(a) Offspring (b) Parents

**Figure 3: Population evolution.**

## Acknowledgments

This work was partially supported by the Conselho Nacional de Pesquisa e Desenvolvimento (CNPq) - Brazil - Finance Codes: [310195/2022-5-PQ2], [UNIVERSAL 4408164/2021-2], PDE [200014/2025-0]. The first author is immensely grateful for the support and love of Riu, Osti, Gyuri, Loona and EPCA.

## References

[1] Tore Hägglund and José Luis Guzmán. 2024. Give us PID controllers and we can control the world. *IFAC-PapersOnLine* 58, 7 (2024), 103–108.
[2] Gilberto Reynoso-Meza, Jesús Carrillo-Ahumada, Victor Henrique Alves Ribeiro, and Tainara Marques. 2022. Multi-objective PID Controller Tuning for Multi-model Control of Nonlinear Systems. *SN Computer Science* 3, 5 (2022), 351.
[3] Gilberto Reynoso-Meza, Javier Sanchis, Xavier Blasco, and Sergio Garcia-Nieto. 2014. Physical programming for preference driven evolutionary multi-objective optimization. *Applied Soft Computing* 24 (2014), 341–362.

Received 20 June 2025; revised —; accepted —