Optimizing Neural Network Performance in Game Playing Using Simulated Annealing and Reinforcement Learning

Henrique Coutinho Layber[1], **Vitor Berger Bonella**[1], **Flávio Miguel Varejão**[1]

[1]Departamento de Informática – Universidade Federal do Espírito Santo (UFES)
Vitória – ES – Brazil

{henrique.layber,vitor.bonella}@edu.ufes.br,flavio.varejao@ufes.br

Abstract. *This paper experiments optimization for Neural Network (NN) parameters for game playing using Simulated Annealing (SA) and Reinforcement Learning (RL). The study focuses on the Dino Game, comparing the performance of the proposed NN method against a baseline Decision Tree method. Experimental results demonstrate that the NN outperforms the Decision Tree, achieving a higher mean score with greater consistency. Statistical tests confirm the performance improvements are statistically significant, indicating the effectiveness of the SA heuristic in optimizing NN parameters.*

1. Introduction

The success of machines in surpassing human performance in games has spurred numerous advancements in Artificial Intelligence (AI), particularly through Neural Networks (NN) [Silver et al. 2016]. NN's ability to handle complex tasks, including game playing, has been explored extensively, often in conjunction with Reinforcement Learning (RL) [Sutton and Barto 2018]. RL enables an agent to learn optimal actions by interacting with its environment and observing outcomes, typically involving a search heuristic and a classifier.

This paper investigates optimizing NN parameters for the Dino Game [Kulkarni et al. 2023] using the Simulated Annealing (SA) heuristic [Burkard and Rendl 1984]. Section 2 presents the proposed method, Section 3 details the experimental setup, Section 4 outlines the results, and Section 5 summarizes the findings and potential future work.

2. Problem Description

This section details the methodology for optimizing the NN for the Dino Game using the Simulated Annealing (SA) heuristic.

2.1. Dino Game

The Dino Game[1] is a runner game where the player controls a dinosaur to avoid obstacles by jumping or crouching. Key variables impacting player actions include game speed, distance to the next object, and object height. Figure 1 illustrates the game frame, highlighting the variables.

[1]Officially available on `chrome://dino`, for Chromium browsers only

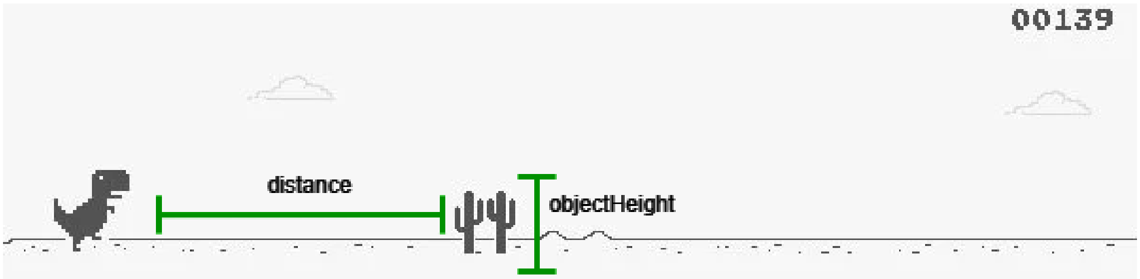

Figure 1. Dino Game Frame

2.2. Neural Network

Neural Networks (NNs) are central to modern AI [McCulloch and Pitts 1943], functioning as function approximators in RL, mapping states to actions based on environmental feedback. The NN used in this study receives three inputs: distance to the next object, object height, and game speed (Figure 2). No biases were included in the network.

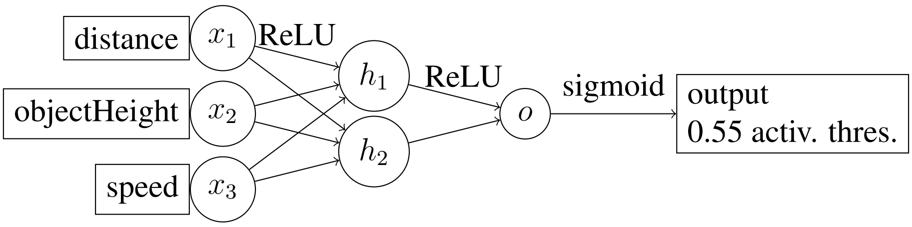

Figure 2. Neural Network Architecture

2.3. Simulated Annealing

SA is a probabilistic technique inspired by the annealing process in metallurgy, useful for finding near-optimal solutions in large search spaces. In this study, SA is applied to optimize NN parameters for decision-making in the Dino Game. The cooling rate, a key meta-parameter, balances exploration and exploitation during optimization, with different cooling schedules like Boltzmann [Kirkpatrick et al. 1983], exponential [Van Laarhoven et al. 1987], and geometric [Černỳ 1985] cooling strategies explored in this paper.

3. Experimental Setup

The experiments compare the proposed method with a baseline Decision Tree optimized using gradient ascent [Boyd and Vandenberghe 2004], both based on the input variables in Section 2.

Python was used to implement the setup[2], with the NN weights optimized using SA over an eight-hour training session. SA's settings included a starting temperature of 200 and geometric cooling with a 0.003 rate. Each solution was tested 15 times to ensure stability, with the final score being the mean minus the standard deviation. To compare the methods, the *paired t-test* and *Wilcoxon test* were used to determine statistical significance. The best solutions were tested 30 times to ensure normality assumptions were met. Figure 3 shows the score evolution across epochs.

[2] Available on GitHub.

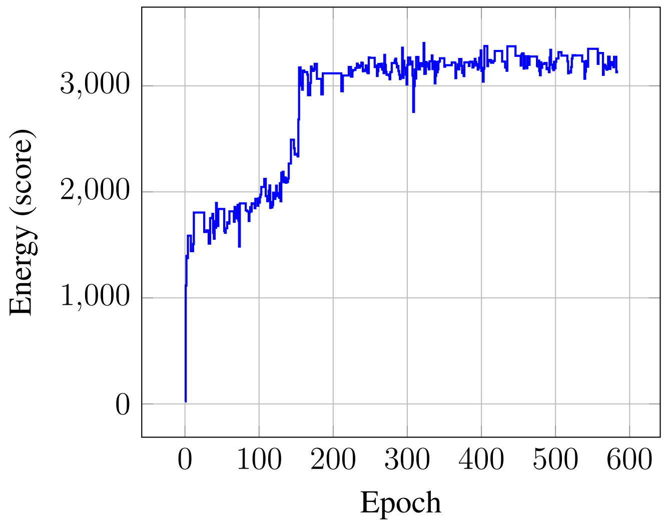

Figure 3. Energy evolution

4. Results and Discussions

The performance of the NN method and the baseline Decision Tree was evaluated based on experimental data. Decision Tree was picked for a baseline for being easily interpretable. Figure 4 compares the score distributions, showing that the NN not only achieves higher median scores but also demonstrates less variability compared to the Decision Tree.

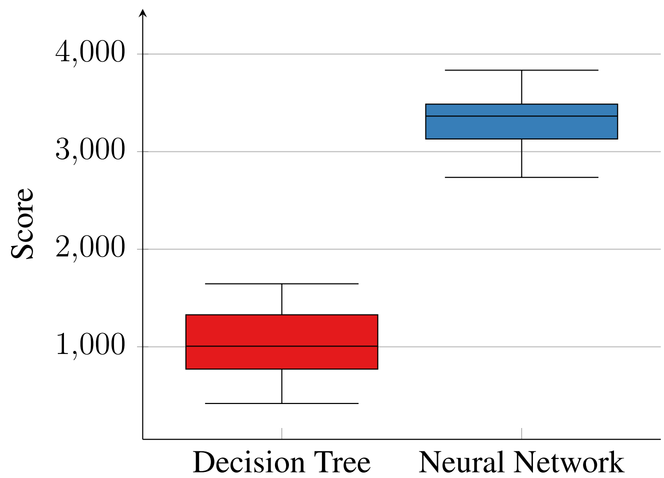

Figure 4. Score boxplots of the methods

The *paired t-test* and *Wilcoxon test* both show p-values near zero, confirming that the performance differences between the NN and the Decision Tree are statistically significant at the 95% confidence level.

Figure 3 highlights two key performance improvements: an initial jump to around 1200 scores, indicating basic action learning, and a more significant increase between epochs 150 and 154, reflecting enhanced timing and performance at higher game speeds: by learning to anticipate and react to obstacles earlier, the agent significantly enhances its overall performance and stability.

5. Conclusion

This paper introduces a novel method for optimizing NN parameters using SA and RL in the context of the Dino Game. Our experiments demonstrate that the NN outperforms a baseline Decision Tree in both mean score and consistency.

Statistical analysis with *paired t-test* and *Wilcoxon test* confirms the NN's performance improvements are significant. This study highlights SA's potential in tuning NN parameters and the effectiveness of NNs in complex game scenarios.

Future work could apply this method to other games and investigate additional optimization techniques, such as different cooling schedules or combining SA with other meta-heuristics. This research adds to the existing body of work on AI in game playing and offers potential avenues for further exploration in AI and Machine Learning.

References

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Burkard, R. and Rendl, F. (1984). A thermodynamically motivated simulation procedure for combinatorial optimization problems. *European Journal of Operational Research*, 17(2):169–174.

Černỳ, V. (1985). Thermodynamical approach to the traveling salesman problem: An efficient simulation algorithm. *Journal of optimization theory and applications*, 45:41–51.

Kirkpatrick, S., Gelatt Jr, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *science*, 220(4598):671–680.

Kulkarni, A., Bapat, P., Kulkarni, T., and Pawar, R. (2023). Review of reinforcement learning in chrome dino game. In *2023 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA)*, pages 1–5.

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Van Laarhoven, P. J., Aarts, E. H., van Laarhoven, P. J., and Aarts, E. H. (1987). *Simulated annealing*. Springer.