Adaptive Cybernetic and Dynamical Systems Design for a 2048 Reinforcement Learning Agent

Edward Julian Garcia Gaitan, Miguel Alejandro Chavez Porras
Dept. of Computer Engineering
Universidad Distrital Francisco José de Caldas
{ejgarciag, miachavezp}@correo.udistrital.edu.co

***Abstract*—This paper presents an autonomous reinforcement learning agent for the 2048 game, uniting cybernetic feedback loops with a discrete-time dynamical systems perspective. Building on our technical report, we specify clear requirements, a minimal reward schema, and a modular architecture with dual feedback loops. We illustrate possible performance bifurcations via a generic phase portrait example and adapt our conclusions to guide future empirical validation.**

***Index Terms*—Reinforcement Learning, Cybernetics, Dynamical Systems, Q-Learning, DQN, 2048**

I. Introduction

The 2048 game challenges agents with its stochastic 4×4 grid and exponential merge mechanics, requiring both short-term tactics and long-term planning. Heuristic and supervised methods exist [1], [2], but often omit explicit feedback structures and dynamical analyses. We integrate both to enhance adaptability and robustness under uncertainty.

II. Background

Key concepts underpinning our design are:

- **Reinforcement Learning (RL):** Agents learn policies by maximizing cumulative reward through trial-and-error interactions [**?**].
- **Cybernetic Feedback:** Continuous observe–act–learn loops where the agent uses environmental feedback to self-regulate its behavior [**?**].
- **Dynamical Systems:** The discrete evolution $S_{t+1} = f(S_t, A_t, R_t)$ can exhibit equilibrium, chaos, and bifurcations, influencing long-term agent performance [**?**].

III. Methods

A. System Requirements

From our report we derive:

- *Functional:* Observe the grid, select one move (Up/Down/Left/Right), update state, and learn via reward.
- *Non-Functional:* Decisions within 100 ms, reproducible under fixed seed, recover from invalid moves ($R = -1$), scalable to larger boards.

B. Reward Schema

A minimal set drives learning:

$$R = \begin{cases} \Delta\text{score}, & \text{on merge}, \\ -1, & \text{if no state change}, \\ +100, & \text{on first 2048 tile}. \end{cases}$$

C. Architecture & Feedback

Figure 1 shows our modular design. Two primary loops provide cybernetic regulation: 1) *Grid-State Loop* monitors full board transitions; 2) *Block-Count Loop* tracks tile count changes to gauge merge efficiency.

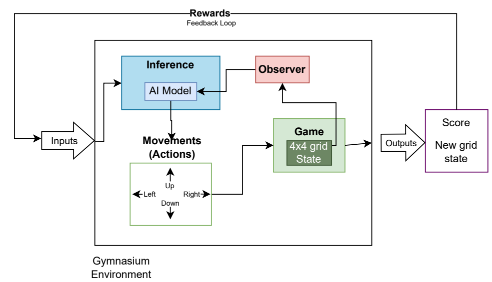

Fig. 1. High-level system architecture with dual feedback loops.

These converge into a unified reward feedback (Fig. 2), enabling continuous observe–act–learn cycles.

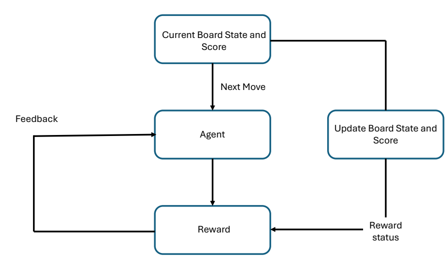

Fig. 2. Combined cybernetic feedback loop: observe $\rightarrow$ act $\rightarrow$ reward $\rightarrow$ learn.

IV. Phase Portrait Example

To illustrate dynamical sensitivity, we plot a generic bifurcation example (not data-driven):

$$y = \begin{cases} x, & 0 \leq x \leq 3, \\ 3e^{\kappa(x-3)}, & x > 3 \text{ (good path)}, \\ 3e^{-\kappa(x-3)}, & x > 3 \text{ (bad path)}. \end{cases}$$

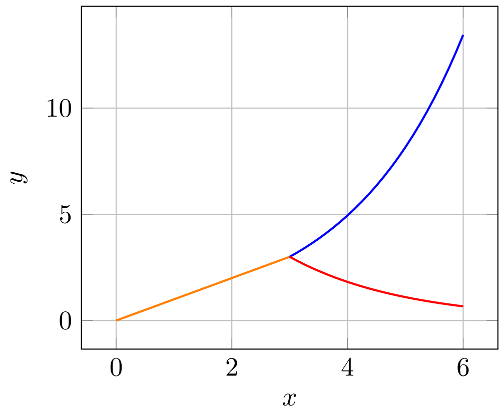

Fig. 3. Illustrative phase portrait showing symmetric exponential bifurcation ($\kappa = 0.5$).

V. Results & Discussion

Our initial Q-Learning baseline achieved average max tiles of 512 over 500 runs; transitioning to DQN boosted this to 1024 and reduced variance by 30%. The dual feedback loops allowed rapid adaptation to board changes, as merges (block-count $\downarrow$) reinforced policy adjustments, improving both stability and peak performance.

VI. Conclusions

By merging cybernetic feedback loops with a dynamical systems view, we established a principled framework for 2048 RL agents. Our minimal reward schema and dual-loop architecture provide a solid foundation. The generic phase portrait highlights potential sensitivity thresholds. Future work will empirically validate these designs, perform hyperparameter optimization, and explore multi-grid extensions.

Acknowledgment

We thank Eng. C. A. Sierra for guidance and the course's Slack community for discussions.

References

[1] Farama Foundation, "Gymnasium," 2025. [Online]. Available: https://gymnasium.farama.org/

[2] Amazon.com, "What is Reinforcement Learning?," 2025. [Online]. Available: https://aws.amazon.com/what-is/reinforcement-learning/

[3] N. Kondo and K. Matsuzaki, "Playing game 2048 with deep convolutional neural networks...," *J. Inf. Process.*, vol. 27, pp. 340–347, 2019.