Proof-of-Attention (PoA): Game-Theoretic and Information-Theoretic Analysis

Part II-B: Evolutionary Stability, Bayesian Incentives, Welfare, and Calibration

OCTA Research

Abstract

Part II-B extends the PoA economic analysis to evolutionary game dynamics, welfare optimization, Bayesian private-type incentives, and calibration of mint–burn policy curves. We characterize evolutionarily stable strategies, welfare alignment conditions, and attention-entropy–based policy modulation. TikZ/PGFPlots figures illustrate replicator flows and net mint curves.

Contents

1 Evolutionary PoA Game

Let $x_j(t)$ be the fraction of players using strategy a^j.

$$x_j(t+1) = x_j(t)\frac{\pi_j(t)}{\bar{\pi}(t)}.$$

Definition 1.1 (ESS)**.** *A strategy surviving all small mutations.*

2 Social Welfare and Policy Alignment

$$W_t = \sum_i U_{i,t}.$$

Mechanism parameters $(g, h, \alpha,)$ are tuned so that welfare-maximizing behavior equilibrium behavior.

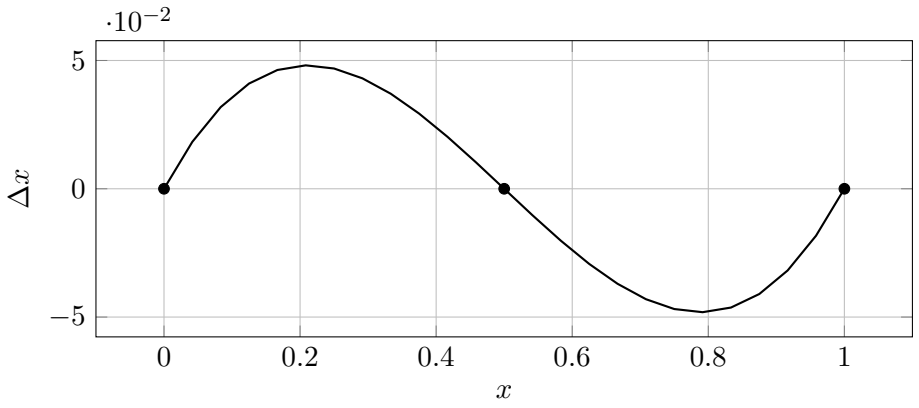

Figure 1: Two-strategy replicator flow.

3 Bayesian PoA Game

Players have private types θ_i affecting costs and preferences.

Definition 3.1 (Bayes–Nash Equilibrium)**.** *Each $\sigma_i(\theta_i)$ maximizes expected utility given beliefs.*

Theorem 3.2 (Informal honesty IC condition)**.** *If humans score higher in and bots trigger stronger fatigue penalties, truthful interaction forms a Bayes–Nash equilibrium for a large class of priors.*

4 Mint–Burn Calibration Example

Let

$$g(x) = \frac{x}{1+x}, \quad h(x) = \eta[x - x_0]_+^2.$$

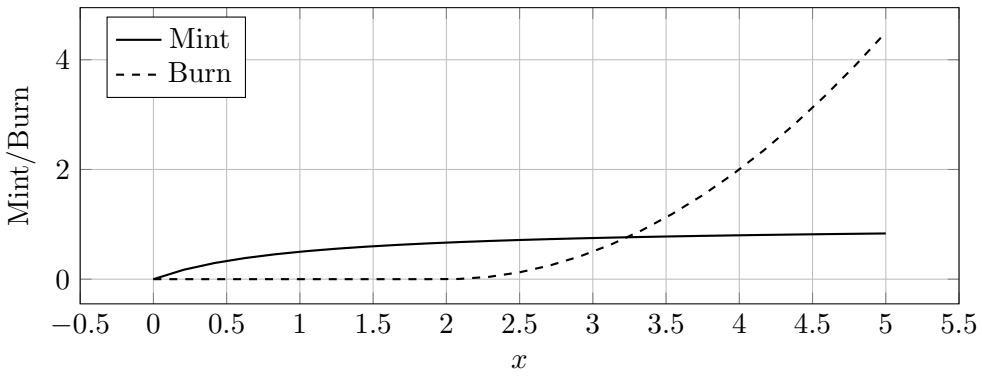

Figure 2: Mint vs burn curves.

5 Conclusion of Part II-B

We characterized:

- ESS and replicator stability,

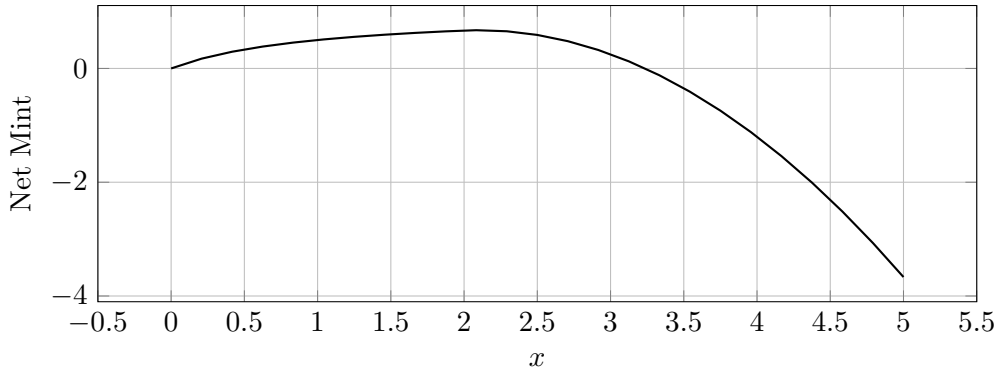

Figure 3: Net mint curve defining stability region.

- welfare alignment targets,
- Bayesian IC constraints,
- calibration of mint–burn controls.

Together with Part II-A, this completes the formal strategic analysis of Proof-of-Attention.