2024 European Control Conference (ECC)
June 25-28, 2024. Stockholm, Sweden

The madness of people: rational learning in feedback-evolving games

Keith Paarporn

Abstract— **The replicator equation in evolutionary game theory describes the change in a population's behaviors over time given suitable incentives. It arises when individuals make decisions using a simple learning process – imitation. A recent emerging framework builds upon this standard model by incorporating game-environment feedback, in which the population's actions affect a shared environment, and in turn, the changing environment shapes incentives for future behaviors. In this paper, we investigate game-environment feedback when individuals instead use a boundedly rational learning rule known as logit learning. We characterize the resulting system's complete set of fixed points and their local stability properties, and how the level of rationality determines overall environmental outcomes in comparison to imitative learning rules. We identify a large parameter space for which logit learning exhibits a wide range of dynamics as the rationality parameter is increased from low to high. Notably, we identify a bifurcation point at which the system exhibits stable limit cycles. When the population is highly rational, the limit cycle collapses and a tragedy of the commons becomes stable.**

I. Introduction

The "Tragedy of the Commons" refers to a scenario in which individuals acting according to their own self-interest leads to the destruction of a shared common resource [1]. The originating example describes a group of cattle herders that share a common pasture land, which becomes overgrazed as each herder allows more of their cows to use it. Indeed, individual incentives are often mis-aligned with collective benefits that could be realized through mutual cooperation. It is relevant to many scenarios: there are economic and personal costs in reducing emissions, quarantining during a pandemic, and conserving resources such as water or gas [2].

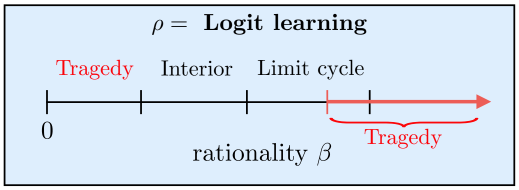
Game theory is a powerful tool that can predict population-level behaviors provided that individuals' incentives can be modeled. Classical formulations predict outcomes when the incentives are static, i.e. they do not change over time. However, actions have consequences on the environment, and a changing environment in turn affects individuals' incentives. For example, when the prevalence of an infectious disease is high, people will tend to stay at home as it becomes more likely to get infected. When the prevalence becomes lower, people will start to resume normal activities – however, this can encourage the spread of new infections [3], [4]. An emerging framework termed "feedback-evolving games" incorporates a dynamic coupling between population-level behaviors and its impact on environmental states [5].

Feedback-evolving games constitutes a flexible framework capable of modeling the coupling between population behaviors and relevant environmental systems, such as social behaviors in epidemics, behaviors in climate change, and consumption of common resources [6]–[9]. Extensive research has characterized many possible dynamics that can emerge [5], [9]–[12]. These feedback-evolving models primarily consider population behaviors that are governed by the replicator dynamics (notable exceptions are [13], [14]). The replicator dynamic arises from simple imitative learning rules at the individual level: an agent changes its action if it observes another agent that is more successful using a different action. The main assumptions underlying imitative learning is that agents do not utilize sophisticated cognitive abilities to make a decision [15]. However, people sometimes make rational choices (i.e. payoff-maximizing), and sometimes make irrational ones (suboptimal) due to noise in their decision-making or available information [16].

In this paper, we depart from the usual replicator model by considering a feedback-evolving game where agents make boundedly rational choices. Instead of imitation, agents are logit learners. The logit rule is parameterized by a rationality parameter $\beta \geq 0$. When $\beta = 0$, agents blindly choose an action uniformly at random. As β becomes higher, agents make a payoff-maximizing choice at higher rates, and a suboptimal choice at lower rates. In the limit of large β, the logit rule converges to a best-response. Logit learning is fundamentally different from imitation, as it requires agents to have access to information about payoffs from all strategies. The well-known algorithm called "log-linear learning" in finite player settings has extensively been studied in regards to its convergence properties in potential games [17]–[19] and networked coordination games [20]–[22].

The primary contribution of this paper is the analysis of the dynamics induced by logit learning for varying levels of the rationality parameter. We focus our study on a parameter regime in which imitative learning is known to lead to a tragedy of the commons as the globally stable outcome– all agents defect and the environment collapses. Thus, our study is also aimed at determining the effectiveness of rational learning in stabilizing more desirable environmental outcomes. Interestingly, the logit system exhibits a variety of dynamics that range from a tragedy of the commons to limit cycles. A summary of our results is depicted in Figure 1.

We provide preliminary background on feedback-evolving games in Section II. The proposed logit dynamics are presented in Section III. Here, we identify the complete set of fixed points of this system and conditions for their stability (Theorem 3.1). We note that the set of fixed points differ from those in the original, imitative system. In Section IV, we more closely analyze properties of an interior fixed point. We identify the rationality level where it undergoes a Hopf bifurcation, which gives rise to stable limit cycles (Theorem 4.1).



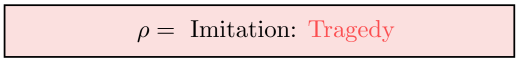


Fig. 1: Contributions: summary of dynamics of logit learning in feedback-evolving games. In increasing order of the population's rationality β, the system exhibits: 1) a tragedy of the commons (TOC), 2) an interior fixed point (sustained non-zero resource), 3) a bifurcation into stable and growing limit cycles, and 4) again a TOC. The system also exhibits bistability between a tragedy and a limit cycle. When agents instead follow imitative protocols, the dynamics always leads to a tragedy.

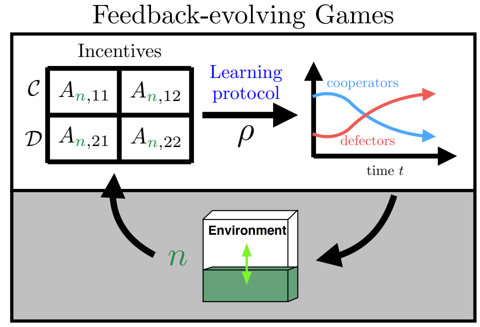


Fig. 2: The feedback-evolving games framework. Individuals' actions affect a shared environment, which shapes incentives for future actions. The agents' decision-making process is specified by the learning protocol ρ.

II. Background: Feedback-Evolving Games

A feedback-evolving game considers a population of agents whose actions have consequences on the abundance of an environmental state or shared resource, $n \in [0,1]$ (Figure 2). At any given time, an agent chooses whether to cooperate ($\mathcal{C}$) or defect ($\mathcal{D}$). The defect action degrades n (e.g. high resource consumption), and the cooperate action contributes to improving n (e.g. restrained consumption). The immediate payoffs to each action are specified by the following 2×2 payoff matrix

$$A_n = n \begin{bmatrix} R_1 & S_1 \\ T_1 & P_1 \end{bmatrix} + (1-n) \begin{bmatrix} R_0 & S_0 \\ T_0 & P_0 \end{bmatrix} \tag{1}$$

where the first row (and column) corresponds to the $\mathcal{C}$ action, and the second row (and column) corresponds to the $\mathcal{D}$ action. The parameters R_0, S_0, T_0, P_0 specify the payoff structure when the environment is in a depleted state ($n = 0$), and R_1, S_1, T_1, P_1 specify the payoff structure when the environment is in a replete state ($n = 1$). Note that the payoff matrix A_n is *environment-dependent*, and thus changes over time. Denoting $x \in [0,1]$ as the fraction of cooperating agents in the population, the payoff to a cooperating and defecting agent is given by

$$\pi_{\mathcal{C}}(x,n) = [A_n[x, 1-x]^\top]_1, \quad \pi_{\mathcal{D}}(x,n) = [A_n[x,1-x]^\top]_2 \tag{2}$$

respectively. We will denote the payoff difference between cooperation and defection as:

$$\begin{aligned} g(x,n) &:= \pi_{\mathcal{C}}(x,n) - \pi_{\mathcal{D}}(x,n) \\ &= axn + bx + cn + d \end{aligned} \tag{3}$$

where we denote

$$\begin{aligned} a &:= \delta_{SP0} - \delta_{RT0} + \delta_{PS1} - \delta_{TR1} \\ b &:= \delta_{RT0} - \delta_{SP0} \\ c &:= -(\delta_{PS1} + \delta_{SP0}) \\ d &:= \delta_{SP0}. \end{aligned} \tag{4}$$

and $\delta_{TR1} = T_1 - R_1$, $\delta_{PS1} = P_1 - S_1$, $\delta_{RT0} = R_0 - T_0$, and $\delta_{SP0} = S_0 - P_0$. The δ constants are referred to as payoff parameters.

Assumption 1. *Defection is the dominant strategy in the 2×2 game that corresponds to the payoff matrix A_1. In particular, $\delta_{TR1} > 0$ and $\delta_{PS1} > 0$.*

The above assumption is widely adopted in the feedback-evolving games literature. It asserts that agents have more incentives to consume resources when they are abundant ($n = 1$).

The agents dynamically update their decisions over time. In the standard analyses of feedback-evolving games, agents are assumed to follow a revision protocol that induces the replicator dynamics. A revision protocol is a description of the behavioral dynamics of agents in the population. It is specified by a function $\rho_{ij}(x,n)$ that gives the rate at which an agent currently adopting strategy i switches to strategy j. For a given revision protocol ρ_{ij}, the *mean population dynamics* describing the change in cooperator fraction over time is generically given by the rate equation

$$\dot{x} = (1-x)\rho_{\mathcal{DC}}(x,n) - x\rho_{\mathcal{CD}}(x,n). \tag{5}$$

Imitative revision protocols induce the replicator dynamics – one such example is the imitative pairwise comparison protocol

$$\rho_{\mathcal{DC}}(x,n) = x[g(x,n)]_+, \quad \rho_{\mathcal{CD}}(x,n) = (1-x)[-g(x,n)]_+ \tag{6}$$

where $[a]_+ = \max\{0, a\}$. A defecting agent will switch to cooperate only if $g(x,n) > 0$, and a cooperating agent will switch to defection only if $g(x,n) < 0$. The overall coupled game-environment system dynamics considered in [5] is then given by

$$\begin{aligned} \dot{x} &= x(1-x)g(x,n) \\ \dot{n} &= \epsilon n(1-n)(\theta x - (1-x)) \end{aligned} \tag{ID}$$

The form of the environmental dynamics $F_n(x,n)$ is referred to as the *tipping point dynamics*, and has been extensively studied in the literature [5], [9]–[11], [13]. The environment does not improve unless a sufficient fraction of the population cooperates. The parameter $\epsilon > 0$ is a time-scale separation constant. We note that the choice of ϵ can be arbitrary, as it does not affect any of the stability analysis in this work.

The cooperators help restore the environment at the rate θ, and defectors degrade n at a unit rate. The state $\boldsymbol{z} = (x, n)$ evolves over the state space $\Gamma := [0,1]^2$. By inspection, one can verify that Γ is forward-invariant with respect to the dynamics (ID).

We will classify two types of fixed points of the feedback-evolving system (ID). A *tragedy of the commons (TOC)* is a fixed point of the form $\boldsymbol{z}_{\rm t} = (x_{\rm t}, 0)$. Such an outcome indicates that the environmental resource has totally collapsed. A *prosperity* fixed point is of the form $(x, 1)$. An *interior fixed point* is one such that $\boldsymbol{z}^* \in \text{int}(\Gamma)$, i.e. $x^*, n^* \in (0,1)$. The imitative system (ID) has four corner fixed points $(0,0), (0,1), (1,0), (1,1)$, and under some parameter regimes, a unique interior fixed point.

The goal of this paper is to characterize how the above system dynamics qualitatively differ when agents follow an alternate revision protocol known as *logit learning*, which endows agents with some degree of rationality. Our comparative analysis will focus on a parameter regime in which the standard imitative dynamics (ID) is known to lead to a tragedy of the commons. Thus, we will make the following assumptions.

Assumption 2. *The replenishment rate $\theta < 1$.*

In words, defection degrades the resource faster than cooperation restores it. Consequently, an irrational population will cause a tragedy of the commons. Moreover, we consider the following condition on payoff parameters in the collapsed state:

Assumption 3. *We will consider payoff parameters $\delta_{SP0} < 0$ and $\delta_{RT0} > -\delta_{SP0}$.*

Under Assumptions 2 and 3, it is established in [5] that the tragedy fixed point $(0,0)$ is globally attracting, and there exists a unique and unstable interior fixed point

$$x_{\text{int}} := \frac{1}{1+\theta}, \quad \bar{n} := \frac{\delta_{RT0} + \theta\delta_{SP0}}{\delta_{RT0} + \delta_{TR1} + \theta(\delta_{SP0} + \delta_{PS1})}. \tag{7}$$

III. Model: Logit Learning

Suppose agents follow a perturbed best-response dynamic called the logit protocol [15]. This is a departure from usual considerations that the agents are imitative learners. The logit revision protocol is given by

$$\begin{aligned}\rho_{\mathcal{C}}(x,n) &= \frac{e^{\beta\pi_{\mathcal{C}}(x,n)}}{e^{\beta\pi_{\mathcal{C}}(x,n)} + e^{\beta\pi_{\mathcal{D}}(x,n)}} \\ \rho_{\mathcal{D}}(x,n) &= \frac{e^{\beta\pi_{\mathcal{D}}(x,n)}}{e^{\beta\pi_{\mathcal{C}}(x,n)} + e^{\beta\pi_{\mathcal{D}}(x,n)}}\end{aligned} \tag{8}$$

where $0 \leq \beta < \infty$ is the rationality parameter of an agent. The logit protocol is fundamentally different from imitative protocols. The protocol $\rho_{\mathcal{C}}$ (resp. $\rho_{\mathcal{D}}$) describes the switch rate to strategy $\mathcal{C}$ ($\mathcal{D}$) for any agent in the population. For low values of β, agents choose their actions uniformly at random, and for high values of β, they select the payoff-maximizing action with a probability close to 1. From (5), the logit protocol induces the mean dynamics

$$\begin{aligned}\dot{x} &= (1-x)\rho_{\mathcal{C}}(x,n) - x\rho_{\mathcal{D}}(x,n) \\ &= (1-x)\rho_{\mathcal{C}}(x,n) - x(1-\rho_{\mathcal{C}}(x,n)) \\ &= \rho_{\mathcal{C}}(x,n) - x\end{aligned} \tag{9}$$

and overall, the coupled game-environment system dynamics are:

$$\boxed{\begin{aligned}\dot{x} &= F_1(x,n) := \frac{e^{\beta g(x,n)}}{1+e^{\beta g(x,n)}} - x \\ \dot{n} &= F_2(x,n) := \epsilon n(1-n)(\theta x - (1-x))\end{aligned}} \tag{LD}$$

The system (LD) will be the main focus of this paper. One can verify that the state space Γ is forward-invariant through an application of Nagumo's Theorem: whenever $x = 0$ or $x = 1$, $\dot{x} > 0$ or $\dot{x} < 0$, respectively. Moreover, when $n = 0$ or $n = 1$, it holds that $\dot{n} = 0$.

A. Characterization of fixed points

We will classify a *logit interior fixed point* as a fixed point of system (LD) that satisfies $\boldsymbol{z}_l^* = (x^*, n^*) \in \text{int}(\Gamma)$. We observe that when $\beta = 0$, every agent chooses an action uniformly at random, and thus in equilibrium, $x = 1/2$. The equilibrium environmental state is then determined by the value of θ: under the assumption $\theta < 1$, a completely irrational population causes a tragedy $n = 0$.

It is important to note that the logit system does not share any of the fixed points as the imitative system (ID). However, we may still classify fixed points as either TOC or interior. The complete set of fixed points of (LD) is characterized in the result below.

Theorem 3.1. *The fixed points of system* (LD) *are characterized as follows.*

1) *Suppose $\beta = 0$. If $\theta \neq 1$, then there are two fixed points $(\frac{1}{2}, 0)$ and $(\frac{1}{2}, 1)$. If $\theta = 1$, then a line of equilibria $(\frac{1}{2}, n)$ for all $n \in [0,1]$ exists.*
2) *Suppose $\beta > 0$. A unique interior fixed point $\boldsymbol{z}_{int}^* = (x_{int}, n_{int})$ exists if and only if*
$$\beta \geq \beta_{int} := \frac{(1+\theta)\log\theta^{-1}}{\delta_{RT0} + \theta\delta_{SP0}} \tag{10}$$
where x_{int} was given in (7) *and n_{int} is given by*
$$n_{int} = \frac{\delta_{RT0} + \theta\delta_{SP0} - \beta^{-1}(1+\theta)\log\frac{1}{\theta}}{\delta_{RT0} + \theta\delta_{SP0} + \delta_{TR1} + \theta\delta_{PS1}} \in (0,1) \tag{11}$$
If $\beta < \beta_{int}$, no interior fixed points exist.
3) *Suppose $\beta > 0$. There exists a $\hat{\beta} > 4/b$ such that:*
   a) *For all $\beta \in (0, \hat{\beta})$, there is a unique TOC fixed point $x_{t3}(\beta)$. It holds that it is strictly increasing in β, $x_{t3}(\beta) > \frac{1}{2}$, $x_{t3}(\beta_{int}) = x_{int}$, and $\lim_{\beta\to\infty} x_{t3}(\beta) = 1$.*
   b) *For all $\beta \geq \hat{\beta}$, there are three TOC fixed points $x_{t1} \leq x_{t2} < \frac{1}{2} < x_{t3}$, where the equality holds if and only if $\beta = \hat{\beta}$. It holds that x_{t1} is strictly decreasing with $\lim_{\beta\to\infty} x_{t1}(\beta) = 0$, and x_{t2} is strictly increasing with $\lim_{\beta\to\infty} x_{t2}(\beta) = \frac{|\delta_{SP0}|}{\delta_{RT0}}$.*

4) *Suppose $\beta > 0$. There exists a unique fixed point of the form $(x^*, 1)$, where $x^* < \frac{1}{2}$ is strictly decreasing in β and $\lim_{\beta\to\infty} x^*(\beta) = 0$.*

The proof is omitted for brevity, but is available in an online version [23]. The environmental level at the interior fixed point n_{int} is monotonically increasing in the rationality $\beta > \beta_{\text{int}}$. At the threshold $\beta = \beta_{\text{int}}$, $n_{\text{int}} = 0$, and as $\beta \to \infty$, the level n_{int} approaches $\bar{n}$, the interior fixed point from the imitative system (7). For identifying TOC fixed points, the fixed point equation (LD) is transcendental, and thus its solutions cannot be expressed generally in closed form. Consequently, the precise value of $\hat{\beta}$ in item 3) cannot be analytically derived.

B. Stability properties of fixed points

To conclude this section, we establish the stability properties of all fixed points except the interior FP, which we will closely investigate in the next section. When $\beta = 0$, the only two fixed points are $(1/2, 0)$ and $(1/2, 1)$. From Assumption 2, we immediately deduce that $(1/2, 0)$ is stable and $(1/2, 1)$ is unstable. Since no interior fixed point exists at $\beta = 0$, we may also conclude that $(1/2, 0)$ is globally attractive by invoking Poincare-Bendixson Theorem (there cannot be any orbits in int Γ).

The following result details the stability properties of TOC and prosperity fixed points when $\beta > 0$.

Proposition 3.1. *Suppose $\beta > 0$.*

1) *A TOC fixed point $(x_t, 0)$ is locally stable if and only if*

$$x_t < x_{int} \text{ and } \beta b x_t(1 - x_t) < 1. \tag{12}$$

For $\beta \geq \beta_{int}$, x_{t3} is unstable. For sufficiently large $\beta \geq \hat{\beta}$, x_{t1} is stable and x_{t2} is unstable.

2) *The fixed point of the form $(x^*, 1)$ is unstable.*

The proof is omitted for brevity, but is provided in the online version [23]. It is interesting to note that item 1 above suggests that for sufficiently high rationality, a TOC is a locally asymptotically stable outcome (x_{t1}). Moreover, this result implies that only TOC and interior fixed points can be stable. We will see in the next section that for intermediate values of β, the system exhibits bifurcations of the interior fixed point n_{int}.

IV. Bifurcations from Logit Learning

In this section, we take the rationality level $\beta > 0$ as a bifurcation parameter of the logit system (LD). We study the stability properties of the interior fixed point as β increases. Notably, we establish a critical value β_h at which it undergoes a Hopf bifurcation. That is, for a neighborhood of values $\beta > \beta_h$, the system exhibits a stable limit cycle around the fixed point $(x_{\text{int}}, n_{\text{int}})$ whose amplitude grows in β. A summary of the system's behavior is illustrated as a bifurcation plot in Figure 3.

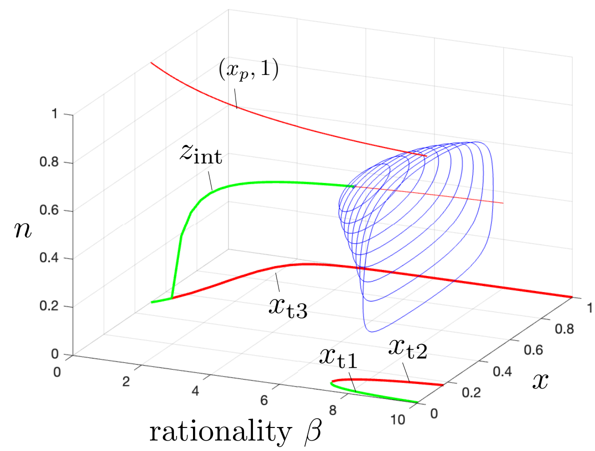


Fig. 3: *Bifurcation plot of the logit system* (LD). Fixed parameters are: $\epsilon = 0.5$, $\delta_{TR1} = 0.5$, $\delta_{PS1} = 0.25$, $\delta_{SP0} = -0.5$, $\delta_{RT0} = 1.5$, $\theta = 0.8$. Thresholds are: $\beta_{\text{int}} = 0.3651$, $\beta_h = 5.6767$. The green lines indicate stable, isolated fixed points. The red lines indicate unstable isolated fixed points. The blue lines are stable limit cycles. They exist only in the interval $\beta \in (\beta_h, \beta_u)$, where $\beta_u \approx 7.84$. For $\beta > \beta_u$, The TOC fixed point x_{t1} appears globally stable.

A. Bifurcation of limit cycles

We first state the Hopf bifurcation theorem below.

Theorem 4.1 (Hopf Bifurcation Theorem (Ch. 3 [24])**.** *Consider a dynamical system $\dot{z} = F(z; \beta)$, where $z \in \mathbb{R}^2$ is the state and $\beta \in \mathbb{R}$ is a bifurcation parameter. Suppose the system has an equilibrium $(z^*; \beta^*)$ (where z^* may depend on β) at which the following properties hold:*

1) *The Jacobian evaluated at $(z^*; \beta^*)$ has a pair of pure imaginary eigenvlaues $\lambda(\beta^*) = \pm\omega i$, where $i := \sqrt{-1}$ and $\omega \in \mathbb{R}$.*
2) $\frac{dRe(\lambda(\beta))}{d\beta}\big|_{\beta^*} \neq 0$.

Then the dynamics undergo a Hopf bifurcation at $(z^*; \beta^*)$, which induces a family of periodic solutions in a sufficiently small neighborhood of $(z^*; \beta^*)$.

The properties of the interior fixed point (11) is summarized in the following main result.

Theorem 4.2. *The interior fixed point $\boldsymbol{z}_{int}(\beta) = (x_{int}(\beta), n_{int}(\beta))$ is locally stable for $\beta \in [\beta_{int}, \beta_h)$, where*

$$\beta_h := (a\bar{n}+b)^{-1}\left(\theta(1+\theta^{-1})^2 + \frac{a(1+\theta)}{D}\log\theta^{-1}\right) \tag{13}$$

and $\bar{n}$ was defined in (7). At the value $\beta = \beta_h$, it undergoes a Hopf bifurcation where its eigenvalues are purely imaginary, and it becomes an unstable focus in a vicinity $\beta > \beta_h$.

Proof. The Jacobian of (LD) evaluated at the interior fixed point is

$$J_{\text{int}} := \begin{bmatrix} \frac{\beta \frac{\partial g}{\partial x}^*}{\theta(1+\theta^{-1})^2} - 1 & \frac{\beta \frac{\partial g}{\partial n}^*}{\theta(1+\theta^{-1})^2} \\ \epsilon n_{\text{int}}(1 - n_{\text{int}})(1 + \theta) & 0 \end{bmatrix} \tag{14}$$

where for compactness, we write $\frac{\partial g}{\partial x}^*$ and $\frac{\partial g}{\partial n}^*$ to represent

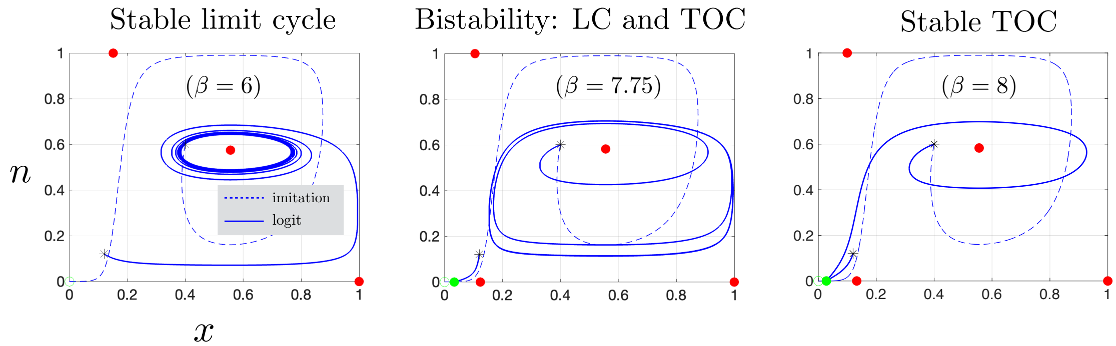


Fig. 4: This figure shows how the system undergoes fundamental changes in its stability properties. To best illustrate this, phase portraits are shown for the selected rationality parameters $\beta = 6, 7.75, 8$. The solid blue lines correspond to the logit dynamics (LD). The dashed blue lines correspond to the standard imitative dynamics (ID). The two asterisks indicate initial conditions. The filled circles are fixed points of the logit system, where red indicates it is unstable and green indicates it is locally stable. The open green circle at $(0, 0)$ is the TOC fixed point from imitative dynamics. Here, we use the same parameter setup as in Figure 3, where the bifurcation value is $\beta_h = 5.67$. (Left) We see the trajectory converges to a stable limit cycle when $\beta = 6$. (Center) The system exhibits bistability when $\beta = 7.75$ – some initial conditions are attracted to x_{t1}, while others are attracted to a stable limit cycle. (Right) For $\beta = 8$, the limit cycle has collapsed and the only stable fixed point of the system is x_{t1}.

the partial derivatives evaluated at $\boldsymbol{z}_{\text{int}}(\beta)$. The trace is

$$\text{tr}J_{\text{int}} = \frac{\beta \frac{\partial g}{\partial x}^*}{\theta(1+\theta^{-1})^2} - 1 \tag{15}$$

and the determinant is

$$\det J_{\text{int}} = -\epsilon\beta \frac{\partial g}{\partial n}^* n_{\text{int}}(1-n_{\text{int}}) \frac{1+\theta}{\theta(1+\theta^{-1})^2}. \tag{16}$$

The fixed point is stable if the real parts of its eigenvalues are negative, which is equivalent to the condition that $\text{tr}J_{\text{int}} < 0$ and $\det J_{\text{int}} > 0$. We have

$$\begin{aligned} \text{tr}J_{\text{int}} < 0 &\iff \beta(an_{\text{int}} + b) < \theta(1+\theta^{-1})^2 \\ &\iff \beta < \beta_h \end{aligned} \tag{17}$$

where the second line follows by re-writing $n_{\text{int}} = \bar{n} - \frac{1+\theta}{D}\log\theta^{-1}$ with $D := \delta_{RT0} + \delta_{TR1} + \theta(\delta_{SP0} + \delta_{PS1}) > 0$, and observing that $a\bar{n} + b = \frac{1+\theta}{D}(\delta_{RT0}\delta_{PS1} - \delta_{SP0}\delta_{TR1}) > 0$. Additionally,

$$\begin{aligned} \det J_{\text{int}} > 0 &\iff -\frac{\partial g}{\partial n}^* = -(ax_{\text{int}} + c) > 0 \\ &\iff -c > a(1+\theta)^{-1} \\ &\iff D > 0. \end{aligned} \tag{18}$$

The sign of $\det J_{\text{int}}$ is determined only from the payoff parameters, and does not depend on β. Thus, $\det J_{\text{int}} > 0$ follows from Assumption 3. This establishes the range of β for which $\boldsymbol{z}_{\text{int}}$ is stable. Its eigenvalues are given by

$$\begin{aligned} \lambda_1(\beta) &:= \frac{\text{tr}J_{\text{int}}}{2} + \sqrt{\left(\frac{\text{tr}J_{\text{int}}}{2}\right)^2 - \det J_{\text{int}}} \\ \lambda_2(\beta) &:= \frac{\text{tr}J_{\text{int}}}{2} - \sqrt{\left(\frac{\text{tr}J_{\text{int}}}{2}\right)^2 - \det J_{\text{int}}} \end{aligned} \tag{19}$$

At the bifurcation point $\beta = \beta_h$, J_{int} has a conjugate pair of purely imaginary eigenvalues $\pm i\omega$ with $\omega = \sqrt{\det J_{\text{int}}} > 0$. Moreover, the rate of change of the eigenvalues' real part is

$$\frac{1}{2}\frac{\partial(\text{tr}J_{\text{int}})}{\partial\beta} = \frac{a\bar{n} + b}{2\theta(1+\theta^{-1})^2} > 0. \tag{20}$$

Indeed, $\text{tr}J_{\text{int}}$ is linearly increasing in β. Therefore, the interior fixed point is an unstable focus (positive real and non-zero imaginary parts) for all values $\beta > \beta_h$ that satisfy $\left(\frac{\text{tr}J_{\text{int}}}{2}\right)^2 < \det J_{\text{int}}$. ■

The Hopf bifurcation at β_h asserts that a family of periodic cycles are guaranteed to appear for a neighborhood of values $\beta > \beta_h$. Whether these periodic cycles are stable depends on the sign of the first Lyapunov coefficient ℓ_1 evaluated at $(\boldsymbol{z}_{\text{int}}; \beta_h)$ (Ch. 3 [24]). If $\ell_1 < 0$, then the bifurcated cycles are stable. The derivation of stability conditions for these cycles will be left for future work. However, we observe through extensive simulations that the bifurcated cycles are stable under the assumed parameter values.

So far, we have established that for $\beta \in [0, \beta_{\text{int}})$, the TOC fixed point x_{t3} is the only stable fixed point (Proposition 3.1). For $\beta \in [\beta_{\text{int}}, \beta_h)$, the interior fixed point $\boldsymbol{z}_{\text{int}}$ is the only stable fixed point (Proposition 3.1 and Theorem 4.1). At $\beta = \beta_h$, it bifurcates into an unstable focus and for a vicinity of values $\beta \geq \beta_h$, a limit cycle encircles $\boldsymbol{z}_{\text{int}}$ (Theorem 4.1). A full bifurcation diagram that summarizes these findings is provided in Figure 3.

B. *Simulations: the high rationality regime*

Under the parameter regime specified by Assumptions 1, 3, and 2, numerical simulations suggest there is another critical value $\beta_u > \beta_h$ for which the limit cycle collapses, and the TOC fixed point x_{t1} becomes globally attractive. Indeed, Proposition 3.1 has established that x_{t1} is stable for sufficiently high β, and it is the only stable fixed point in the system.

Simulations of system trajectories in the phase space are depicted in Figure 4. In particular, we note that the system can exhibit bistability (center portrait) between the limit cycle and x_{t1}: initial conditions close to $(x_{t1}, 0)$ will converge to the TOC, and other conditions will converge to the stable limit cycle. Since the interior fixed point does not disappear for high β, we conjecture that the limit cycle collapses at some value β_u for which its ω-limit set touches the basin of attraction of $(x_{t1}, 0)$. In the example of Figure 4, β_u belongs to the interval $(7.75, 8)$. Such an analysis is left for future work.

V. Conclusion and Future Work

In this paper, we formulated a feedback-evolving system where agents in a population follow a logit revision protocol, which is a fundamentally different learning rule to the usual imitation dynamics. We analyzed the resulting dynamical outcomes as a function of the rationality parameter $\beta \geq 0$. In increasing order of β, we identified interval ranges for which the system exhibits 1) a tragedy of the commons (low rationality), 2) a stable interior fixed point, 3) stable limit cycles, and 4) again, a tragedy of the commons (high rationality). Counter-intuitively, high rationality leads to a collapsed environment, whereas moderate levels of rationality can lead to sustainable outcomes. Our analysis of the logit system holds in a parameter regime where imitative learning leads to a tragedy of the commons. These results demonstrate that boundedly rational behaviors can induce a wide variety of environmental outcomes.

Future work will involve analyzing global stability properties of the system. Additionally, a complete analysis of the Lyapunov coefficient is needed to establish stability of the observed limit cycles, and the precise value of β at which the limit cycle dissipates into the TOC outcome is yet to be established. The application of control strategies, e.g. incentivization of cooperation, to control global outcomes will also be studied.

References

[1] G. Hardin, "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[2] E. Ostrom, *Governing the commons: The evolution of institutions for collective action*. Cambridge university press, 1990.

[3] S. Funk, M. Salathé, and V. A. Jansen, "Modelling the influence of human behaviour on the spread of infectious diseases: a review," *Journal of the Royal Society Interface*, vol. 7, no. 50, pp. 1247–1256, 2010.

[4] J. S. Weitz, S. W. Park, C. Eksin, and J. Dushoff, "Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations," *Proceedings of the National Academy of Sciences*, vol. 117, no. 51, pp. 32 764–32 771, 2020.

[5] J. S. Weitz, C. Eksin, K. Paarporn, S. P. Brown, and W. C. Ratcliff, "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7518–E7525, 2016.

[6] A. Satapathi, N. K. Dhar, A. R. Hota, and V. Srivastava, "Coupled evolutionary behavioral and disease dynamics under reinfection risk," *IEEE Transactions on Control of Network Systems*, 2023.

[7] H. Khazaei, K. Paarporn, A. Garcia, and C. Eksin, "Disease spread coupled with evolutionary social distancing dynamics can lead to growing oscillations," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 4280–4286.

[8] M. Cao and A. Morse, "Modeling the co-evolution of climate impact and population behavior: A mean-field analysis," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 7381–7386, 2023.

[9] A. R. Tilman, J. B. Plotkin, and E. Akçay, "Evolutionary games with environmental feedbacks," *Nature communications*, vol. 11, no. 1, p. 915, 2020.

[10] L. Gong, W. Yao, J. Gao, and M. Cao, "Limit cycles analysis and control of evolutionary game dynamics with environmental feedback," *Automatica*, vol. 145, p. 110536, 2022.

[11] L. Stella, W. Baar, and D. Bauso, "Lower network degrees promote cooperation in the prisoner's dilemma with environmental feedback," *IEEE Control Systems Letters*, vol. 6, pp. 2725–2730, 2022.

[12] K. Paarporn and C. Eksin, "Sis epidemics coupled with evolutionary social distancing dynamics," in *2023 American Control Conference (ACC)*, 2023, pp. 4308–4313.

[13] L. Stella and D. Bauso, "The impact of irrational behaviors in the optional prisoner's dilemma with game-environment feedback," *International Journal of Robust and Nonlinear Control*, 2021.

[14] M. R. Arefin and J. Tanimoto, "Imitation and aspiration dynamics bring different evolutionary outcomes in feedback-evolving games," *Proceedings of the Royal Society A*, vol. 477, no. 2251, p. 20210240, 2021.

[15] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT Press, 2010.

[16] L. E. Blume, "How noise matters," *Games and Economic Behavior*, vol. 44, no. 2, pp. 251–271, 2003.

[17] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.

[18] T. Tatarenko, "Proving convergence of log-linear learning in potential games," in *2014 American Control Conference*. IEEE, 2014, pp. 972–977.

[19] L. E. Blume, "The statistical mechanics of best-response strategy revision," *Games and economic behavior*, vol. 11, no. 2, pp. 111–145, 1995.

[20] V. Auletta, D. Ferraioli, F. Pasquale, and G. Persiano, "Metastability of logit dynamics for coordination games," in *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete algorithms*. SIAM, 2012, pp. 1006–1024.

[21] L. Arditti, G. Como, F. Fagnani, and M. Vanelli, "Equilibria and learning dynamics in mixed network coordination/anti-coordination games," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 4982–4987.

[22] Y. Zhang and M. M. Vasconcelos, "Rationality and connectivity in stochastic learning for networked coordination games," 2023.

[23] K. Paarporn, "The madness of people: rational learning in feedback-evolving games," *arXiv preprint arXiv:2311.02745*, 2023.

[24] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. Springer Science & Business Media, 2013, vol. 42.