Persistent Patterns: Multi-Agent Learning beyond Equilibrium and Utility

Georgios Piliouras⋆, Carlos Nieto-Granda†, Henrik I. Christensen†, Jeff S. Shamma⋆
School of Electrical and Computer Engineering⋆ Institute for Robotics & Intelligent Machines†
Georgia Institute of Technology
{georgios, carlos.nieto, hic, shamma}@gatech.edu

ABSTRACT

We propose an analytic framework for multi-agent learning that, unlike standard approaches, is not connected to convergence to an equilibrium concept nor to payoff guarantees for the agents. We view multi-agent systems as reservoirs that allow for the long term survival of rich spatiotemporal correlations (*i.e.*, patterns) amongst the agents' behaviors. Our aim is to develop abstractions that allow us to capture details about the possible limit behaviors of such systems.

Our approach is based on the contrast between weakly and strongly persistent properties. Informally, a property is weakly persistent if for each starting point there exist limit points that satisfy it. A property is strongly persistent if it is satisfied by all limit points. In the case of non-converging dynamics the set of weakly persistent properties can be significantly richer than that of the strongly persistent properties reflecting topological properties of the system limit sets in a concise and algorithmically tractable manner.

Categories and Subject Descriptors

F.2 [**Analysis of Algorithms**]; J.4 [**Social and Behavior Sciences**]

General Terms

Algorithms, Economics, Theory

Keywords

Multi-Agent Systems, Game Theory, Replicator Dynamics, Information Theory, Topology, Dynamical Systems

1. INTRODUCTION

If multi-agent learning is the answer then what is the question? This inquiry was famously posed in [23], where Shoham, Powers and Grenager initiated a lively discussion joined by AI researchers, game theorists and engineers alike in regards to clarifying the agenda of MAL research [27].

With few notable exceptions [15, 24], this discussion in regards to the scope, goals, and evaluation criteria in multi-agent learning is framed in purely game theoretic terms. Specifically, despite the diverse backgrounds of the authors one would be pressured to identify evaluation criteria of a multi-agent system that do not translate either to some equilibrium notion or to utility guarantees for the agents. Stone [24] voices explicitly such concerns against an overly narrow focus on the game-theoretic framework. Focusing on the example of robotic soccer, he argues that in the case of complex multi-agent environments and domains this predominant game-theoretic perspective does not seem to suffice and one needs a broader perspective. However, no explicit proposal for a novel agenda is put forward. To the contrary, the paper concludes with posing this question as a future challenge to be addressed.

In attempting to provide some insights in this direction, it is useful to remind ourselves that this, undeniably fruitful, relationship between game theory and multi-agent learning has been one-sided from its onset. The motivating goal of the work of Brown and Robinson [6, 19] on fictitious play was to provide validation and computational handles for game theoretic concepts and specifically responding to von Neumann's seminal work on mixed Nash equilibria in zero-sum games [28, 29]. This point of view, with MAL analysis trailing game-theoretic considerations, is based on such solid foundations that it becomes a non-trivial mental exercise to think and argue theoretically about multi-agent systems in a language that is devoid of the notions of equilibrium and utility. How could one accomplish that and why should one care to in the first place?

In order to understand this question let's try to ponder the following: Are the notions of Nash equilibrium and utility in some sense atomic concepts in multi-agent systems, or do they provide a palpable handle on something even more basic? Arguably, *the main desirable characteristic of any system is that it is susceptible to our understanding.* Namely, we aim to identify persistent patterns of actionable data that would allow us to make accurate predictions about future states of the system. Equilibria and utility naturally reside in this framework. The notion of equilibrium, encoding stability, allows for the persistence of static patterns. Utility is the most prevalent signal in economic systems. However, as we will argue, the notion of persistent patterns is significantly richer and allows for important distinctions that would have otherwise been impossible.

Our goal in this work is exactly to explore the notions of patterns and persistence and examine to what extent these ideas can be applied towards a more structured study of dynamics (and especially disequilibrium behavior) in multi-agent systems.

Appears in: *Alessio Lomuscio, Paul Scerri, Ana Bazzan, and Michael Huhns (eds.), Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014), May 5-9, 2014, Paris, France.*

2. FEATURE, PROPERTY AND PATTERN

Generally, let Σ denote a system with a state space $\mathcal{S}$ whose temporal evolution is captured by a flow $\Phi : \mathcal{S} \times J \to \mathcal{S}$, where J is a timeset, that may correspond either to a continuum or a discrete set of time units (*e.g.*, $J = \mathbb{R}, \mathbb{Z}_+$). We define an (observable) feature of Σ as a map $\mathcal{F} : \mathcal{S} \to O$ from $\mathcal{S}$ to (a possibly different) observation space O[1].

In the case of evolutionary dynamics applied to normal form games, where $\mathcal{S}$ corresponds to the product of mixed strategies of the agents, typical examples of observable features are the (mixed) strategy or the (expected) utility of an agent as well as measures of social welfare such as the sum of utilities. However, one can easily come up with other interesting features, such as the median utility, the max-min utility, or diversity indices such as the Herfindahl-Hirschman index. In the case of an ecosystem, where $\mathcal{S}$ captures the densities of each subspecies, an interesting property might project the density of the species that is closest to extinction.

So far we have only considered features that depend on the current state of the system. One can design more involved features that depend on higher order statistics of the system history. An example of such a feature could be the time-average social welfare of a system trajectory. We can capture such features that depend on higher order statistics of the system as standard features of an extended system Σ^*, whose state $\mathcal{S}^*$ keeps track of those statistics as well. Hence, in order to create a comprehensive theory of system feature regularities it suffices to examine properties that depend on the current system state.

The temporal evolution of system Σ induces trajectories on the observation space $\Psi = \mathcal{F} \circ \Phi : \mathcal{S} \times J \to O$ that encode all possible systematic interactions between the system and the $\mathcal{F}$-observer. A systematic analysis of feature $\mathcal{F}$ in Σ implies identifying regularities over all possible trajectories of Ψ. Given feature $\mathcal{F} : \mathcal{S} \to O$, we will denote as property $\Gamma \subset O$ a subset[2] thats encodes a "target" parameter range for that specific feature.

Examples of properties for different features have as follows: If the feature $\mathcal{F}$ is the identity function (*i.e.*, the feature space is the set of all mixed strategy profiles) then a usual property is the set of Nash equilibria of the game, or if the feature is the social welfare (*i.e.*, sum of utilities) then an interesting property is the set of outcomes whose social welfare is within a multiplicative constant of the optimum.

We can also define time-space properties, *e.g.*, subsets of $J \times O$. We call such special properties that capture the arrow of time, patterns. Patterns allow us to capture involved temporal phenomena (*e.g.*, periodic phenomena).

3. PERSISTENCE

The notions of (weak, strong) persistence that we explore here are inspired by more restricted notions of population persistence developed within the field of mathematical ecology ([10, 11] and references therein). The notion of strong persistence, which is analogous to the notion of safety [14, 2] in distributed systems, has recently been used in analyzing optimization questions in dynamical systems [17]. We extend and contrast those techniques with those used to detect and analyze weakly persistent properties, which are analogous to liveness[3]. This combined perspective allows us to capture detailed properties of the topology of multi-agent learning dynamics.

A good starting point for an intuitive understanding of these notions are the ecological inquiries that inspired them. Suppose that we are monitoring an ecosystem by measuring the population sizes of each species. Naturally, we are interested in the health of the ecosystem and a high priority concern is that all species survive in the long run. There exist at least two distinct ways of encoding such guarantees. Informally, strong persistence argues that the population of each species should consistently stay far away from its extinction threshold. Weak persistence allows species to teeter on the brink of extinction infinitely often as long as they always bounce back to a healthy population size. We extend this approach to more general features and properties.

Weak Persistence

Definition 1. *Given feature $\mathcal{F} : \mathcal{S} \to O$ then property $\Gamma_{Weak} \subset O$ is weakly persistent for feature $\mathcal{F}$ if for all initial conditions $x \in \mathcal{S}$,*

$$\limsup_{t\to\infty} dist(\mathcal{F}\big(\Phi(x,t)\big), O \backslash \Gamma_{Weak}) > 0.$$

Furthermore, if $\exists\ \epsilon > 0$ such that $\forall x \in \mathcal{S}$,

$$\limsup_{t\to\infty} dist(\mathcal{F}\big(\Phi(x,t)\big), O \backslash \Gamma_{Weak}) > \epsilon$$

then property Γ_{Weak} is uniformly weakly persistent.

This notion encodes recurring system regularities that despite possibly experiencing lapses of extinction will persistently be revived. Such properties persist in the long run without necessarily being satisfied by all limit points. That is, regardless of the starting state of the system, even if we start from states that do not satisfy a weakly persistent property, such properties will eventually become true for the system and although the system may move away from such states it will always come back to them (see figure 1). The definition of uniform weak persistence is even stronger and essentially requires that this infinitely recurrent property is preserved even in a noisy environment that does not allow for measurements of arbitrary small error.

On the other hand, there exist system properties (*e.g.*, an adequate level of white blood cells) whose satisfaction is a functional imperative for the system. In such cases, we would like to satisfy a significantly stronger persistence guarantee.

(Strong) Persistence

Definition 2. *Given feature $\mathcal{F} : \mathcal{S} \to O$ then property $\Gamma_{Strong} \subset O$ is (strongly) persistent for feature $\mathcal{F}$ if for all initial conditions $x \in \mathcal{S}$,*

$$\liminf_{t\to\infty} dist(\mathcal{F}\big(\Phi(x,t)\big), O \backslash \Gamma_{Strong}) > 0.$$

Furthermore, if $\exists\ \epsilon > 0$ such that $\forall x \in \mathcal{S}$,

$$\liminf_{t\to\infty} dist(\mathcal{F}\big(\Phi(x,t)\big), O \backslash \Gamma_{Strong}) > \epsilon$$

then property Γ_{Strong} is uniformly persistent.

[1] We only consider features that are continuous functions of the (current) state of the system.

[2] For simplicity we assume that O is a compact metric space.

[3] Informally, a safety property stipulates that "bad thing" do not happen during the execution of a program and a liveness property stipulates that "good things" do happen (eventually)[14].

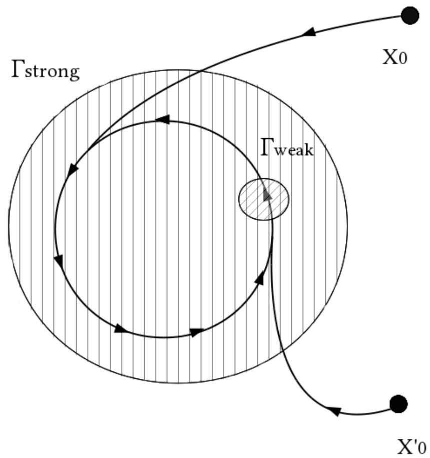

Figure 1: Examples of persistent properties

This notion encodes self-enforcing system regularities. That is, regardless of the starting state of the system, even if we start from states that do not satisfy a persistent property, such properties will eventually become true for the system and persist being true for all time. Any (strongly) persistent property is also weakly persistent.

Why study persistent properties?

Under the assumption of convergence of our system to a fixed point, the clearly distinct notions of weak persistence and (strong) persistence collapse. At a first glance, this implication might actually seem as an advantageous aftereffect of convergence. After all, if a property is desirable enough to be supported in a weakly persistent manner, a system that offers it in a (strongly) persistent manner appears even more desirable. However, in this manner we are restricting the capabilities of our systems in terms of satisfying other properties. Specifically, a system is, at least in theory, able to support in a weakly persistent manner two contradicting property ranges $\Gamma_0, \Gamma_1 \in O, \Gamma_0 \cap \Gamma_1 = \emptyset$ of the same feature $\mathcal{F}$. On the other hand, if one of the property ranges is supported in a persistent manner then the other one cannot be supported, even weakly. So, contrary to the entrenched view of equilibrium convergence as a universally desirable property of systems, *perpetual disequilibrium might be a necessary design specification* for some applications. This framework provides abstractions for capturing tightly the limit behavior of such systems and enables distinctions that would have been impossible if we focused solely on their equilibria.

4. PRELIMINARIES

We combine ideas from (evolutionary) game theory and dynamical systems theory. The presentation here follows in the lines of [17]. Our aim is to provide a concise repository of some key concepts and tools. Several references are included for a more thorough treatment of this material.

4.1 Graphical games

A graphical (polymatrix) game is defined via an undirected graph $G = (V, E)$ where V corresponds to the set of agents of the game and where every edge corresponds to a bimatrix game between its two endpoints/agents. We denote by S_i the set of strategies of agent i. We denote the bimatrix game on edge $(i, k) \in E$ via a pair of payoff matrices: $A^{i,k}$ of dimension $|S_i| \times |S_k|$ and $A^{k,i}$ of dimension $|S_k| \times |S_i|$. Let $s \in \times_i S_i$ be a strategy profile of the game. We denote by $s_i \in S_i$ the strategy of agent i and by $s_{-i} \in \times_{j \in V \setminus i} S_j$ the strategies of the other agents. The payoff of agent $i \in V$ in strategy profile s is equal to the sum of the payoffs that agent i receives from all the bimatrix games she participates in, *i.e.*, $u_i(s) = \sum_{(i,k)\in E} A^{i,k}_{s_i,s_k}$.

A separable zero-sum multiplayer game (zero-sum graphical game) [7] is a graphical polymatrix game in which the sum of all agent payoffs is always zero ($\forall s \in \times_i S_i, \sum_i u_i(s) = 0$). There exists [7] a (polynomial-time computable) payoff preserving transformation from every separable zero-sum multiplayer game to a pairwise constant-sum polymatrix game (*i.e.*, a graphical polymatrix game such that for each $i, k \in V : A^{k,i} = c_{\{k,i\}}\mathbf{1} - \left(A^{i,k}\right)^{\mathrm{T}}$ and $\mathbf{1}$ the all-one matrix).

We will also consider affine transformations of separable zero-sum games. If there exists a separable zero-sum multiplayer game GG and constants $a_i > 0$ and $b_i \in \mathbb{R}$ for each agent i such that $u^{GG}(s) = a_i u_i^G(s) + b_i$ for each outcome $s \in S$ we call such game as $(\vec{a}, \vec{b})$-zero-sum multiplayer game. Affine transformations do not affect the structure of equilibria. They affect, however, the shape and properties of multiagent learning trajectories. A randomized strategy x for agent i lies on the simplex $\Delta(S_i) = \{p \in \mathbb{R}_+^{|S_i|} : \sum_i p_i = 1\}$. Such a strategy x is said to be fully mixed if it lies in the interior of the simplex, *i.e.*, if $x_i > 0$ for all strategies $i \in S_i$.

4.2 Replicator Dynamics

The replicator equation [25, 22] is amongst the most well studied dynamics in evolutionary game theory [10, 12, 17, 13, 20, 21]. It is defined as follows:

$$\dot{x}_i \triangleq \frac{dx_i(t)}{dt} = x_i[u_i(x) - \hat{u}(x)], \quad \hat{u}(x) = \sum_{i=1}^{n} x_i u_i(x)$$

where x_i is the proportion of type i in the population, $x = (x_1, \ldots, x_m)$ is the vector of the distribution of types in the population, $u_i(x)$ is the fitness of type i, and $\hat{u}(x)$ is the average population fitness. The state vector x can also be interpreted as a randomized strategy of an adaptive agent that learns to optimize over its m possible actions given an online stream of payoff vectors. For this reason, it can be applied in games. An interior point of the state space is a fixed point for the replicator if and only if it is a fully mixed Nash equilibrium of the game. The interior (the boundary) of the state space $\times_i \Delta(S_i)$ are invariants for the replicator. We typically analyze the replicator from a generic interior point, since points of the boundary can be captured as interior points of lower dimensional systems. Summing all this up, our model is captured by the following system:

$$\dot{x}_{iR} = x_{iR}\big(u^i(R) - \sum_{R' \in S_i} x_{iR'} u^i(R')\big)$$

for each $i \in N$, $R \in S_i$, where $u^i(R) = \mathrm{E}_{s_{-i} \sim x_{-i}} u_i(R, s_{-i})$.

4.3 Topology of dynamical systems

Our treatment follows that of [30], the standard text in evolutionary game theory, which itself borrows material from the classic book by Bhatia and Szegö [5]. Since our state space is compact and the replicator vector field is Lipschitz-continuous, we can present the unique solution of our ordinary differential equation as a continuous map $\Phi : \mathcal{S}\times\mathbb{R} \rightarrow \mathcal{S}$ called flow of the system. Fixing starting point $x \in \mathcal{S}$ defines a function of time which captures the trajectory (orbit, solution path) of the system with the given starting point. This corresponds to the graph of $\Phi(x, \cdot) : \mathbb{R} \rightarrow \mathcal{S}$, *i.e.*, the set $\{(t, y) : y = \Phi(x, t) \text{ for some } t \in \mathbb{R}\}$.

If the starting point x does not correspond to an equilibrium then we wish to capture the asymptotic behavior of the system (informally the limit of $\Phi(x, t)$ when t goes to infinity). Typically, however, such functions do not exhibit a unique limit point so instead we study the set of limits of all possible convergent subsequences. Formally, given a dynamical system $(\mathbb{R}, \mathcal{S}, \Phi)$ with flow $\Phi : \mathcal{S} \times \mathbb{R} \rightarrow \mathcal{S}$ and a starting point $x \in \mathcal{S}$, we call point $y \in \mathcal{S}$ an ω-limit point of the orbit through x if there exists a sequence $(t_n)_{n\in\mathbb{N}} \in \mathbb{R}$ such that $\lim_{n\rightarrow\infty} t_n = \infty$, $\lim_{n\rightarrow\infty} \Phi(x, t_n) = y$. Alternatively the ω-limit set can be defined as: $\omega_\Phi(x) = \cap_t \overline{\cup_{\tau\geq t}\Phi(x, \tau)}$.

We denote the boundary of a set S as bd(S) and the interior of S as $\text{int}(S)$. In the case of the replicator dynamics where the state space S corresponds to a product of agent (mixed) strategies we will denote by $\Phi_i(x, t)$ the projection of the state on the simplex of mixed strategies of agent i.

Liouville's Formula

Liouville's formula can be applied to any system of autonomous differential equations with a continuously differentiable vector field ξ on an open domain of $\mathcal{S} \subset \mathbb{R}^k$. The divergence of ξ at $x \in \mathcal{S}$ is defined as the trace of the corresponding Jacobian at x, *i.e.*, $\text{div}[\xi(x)] = \sum_{i=1}^{k} \frac{\partial \xi_i}{\partial x_i}(x)$. Since divergence is a continuous function we can compute its integral over measurable sets $A \subset \mathcal{S}$. Given any such set A, let $A(t) = \{\Phi(x_0, t) : x_0 \in A\}$ be the image of A under map Φ at time t. $A(t)$ is measurable and is volume is $\text{vol}[A(t)] = \int_{A(t)} dx$. Liouville's formula states that the time derivative of the volume $A(t)$ exists and is equal to the integral of the divergence over $A(t)$: $\frac{d}{dt}[A(t)] = \int_{A(t)} \text{div}[\xi(x)]dx$.

A vector field is called divergence free if its divergence is zero everywhere. Liouville's formula trivially implies that volume is preserved in such flows.

Poincaré's recurrence theorem

Poincaré [18] proved that in certain systems almost all trajectories return arbitrarily close to their initial position infinitely often.

THEOREM 1. *[18, 3] If a flow preserves volume and has only bounded orbits then for each open set there exist orbits that intersect the set infinitely often.*

Poincaré-Bendixson theorem

The Poincaré-Bendixson theorem allows us to prove the existence of limit cycles[4] in two dimensional systems. The main idea is to find a trapping region, *i.e.*, a region from which trajectories cannot escape. If a trajectory enters and does not leave such a closed and bounded region of the state space that contains no equilibria then this trajectory must approach a periodic orbit as time goes to infinity. Formally, we have:

[4]A periodic orbit is called a limit cycle if it is the ω-limit set of some point not on the periodic orbit.

THEOREM 2. *[4, 26] Given a differentiable real dynamical system defined on an open subset of the plane, then every non-empty compact ω-limit set of an orbit, which contains only finitely many fixed points, is either a fixed point, a periodic orbit, or a connected set composed of a finite number of fixed points together with homoclinic and heteroclinic orbits connecting these.*

Homeomorphisms and Conjugacy of Flows

A function f between two topological spaces is called a *homeomorphism* if it has the following properties: f is a bijection, f is continuous, and f has a continuous inverse. A function f between two topological spaces is called a *diffeomorphism* if it has the following properties: f is a bijection, f is continuously differentiable, and f has a continuously differentiable inverse. Two flows $\Phi^t : A \rightarrow A$ and $\Psi^t : B \rightarrow B$ are conjugate if there exists a homeomorphism $g : A \rightarrow B$ such that for each $x \in A$ and $t \in \mathbb{R}$: $g(\Phi^t(x)) = \Psi^t(g(x))$. Furthermore, two flows $\Phi^t : A \rightarrow A$ and $\Psi^t : B \rightarrow B$ are *diffeomorhpic* if there exists a diffeomorphism $g : A \rightarrow B$ such that for each $x \in A$ and $t \in \mathbb{R}$ $g(\Phi^t(x)) = \Psi^t(g(x))$. If two flows are diffeomorphic then their vector fields are related by the derivative of the conjugacy. That is, we get precisely the same result that we would have obtained if we simply transformed the coordinates in their differential equations [16].

4.4 Information Theory

Entropy is a measure of the uncertainty of a random variable and captures the expected information value from a measurement of the random variable. The entropy H of a discrete random variable X with possible values $\{1, \ldots, n\}$ and probability mass function $p(X)$ is defined as $H(X) = -\sum_{i=1}^{n} p(i) \ln p(i)$. Given two probability distributions p and q of a discrete random variable their *K-L divergence* (relative entropy) is defined as $D_{\text{KL}}(p\|q) = \sum_i \ln\left(\frac{p(i)}{q(i)}\right) p(i)$. A closely related concept is that of the *cross entropy* between two distributions, which measures the average number of bits needed to identify an event from a set of possibilities, if a coding scheme is used based on a given probability distribution q, rather than the "true" distribution p. Formally, the cross entropy for two distributions p and q is equal to $H(p, q) = -\sum_{i=1}^{n} p(i) \ln q(i) = \text{H}(p) + D_{\text{KL}}(p\|q)$. For more details the reader should refer to the standard text by Cover and Thomas [8].

5. ANALYSIS

In this section, we will explore the space of persistent properties for replicator dynamics in graphical polymatrix games. As a default, Φ will denote the flow of the replicator dynamic when applied to an (arbitrary) graphical game and we will focus on the case where the observation function $\mathcal{F}$ is the identity function and $\mathcal{S}$ corresponds to the set of fully mixed strategy profiles.

5.1 Strong Persistence

The analysis of strong persistent properties builds upon techniques developed in [17] for separable zero-sum multi-agent games. We start with a general, negative result that holds for all graphical games.

THEOREM 3. *Let Φ denote the flow of the replicator dynamic when applied to a graphical game and let the observation function $\mathcal{F}$ be the identity function then flow Φ has no uniformly (strongly) persistent property $\Gamma \subset \mathcal{S}$.*

PROOF. The system defined by applying replicator on the interior of state space, can be transformed to a divergence free system on $(-\infty, +\infty)^{\sum_i(|S_i|-1)}$ via the following invertible smooth map $z_{iR} = \ln(x_{iR}/x_{i0})$, where $0 \in S_i$ a specific (but arbitrarily chosen) strategy of agent[5] i. This map $g : \times_i \text{int}(\Delta(S_i)) \to \mathbb{R}^{\sum_i(|S_i|-1)}$ is clearly a homeomorphism[6]. Hence, we can a establish a conjugacy between the replicator system (restricted to the interior of state space) and a system on $(-\infty, +\infty)^{\sum_i(|S_i|-1)}$ where:

$$\frac{d\left(\frac{x_{iR}}{x_{i0}}\right)}{dt} \quad = \quad \frac{\dot{x}_{iR}x_{i0} - \dot{x}_{i0}x_{iR}}{x_{i0}^2} = \frac{x_{iR}}{x_{i0}}\left(u^i(R) - u^i(0)\right).$$

This implies that $\dot{z_{iR}} = \frac{d\left(\ln \frac{x_{iR}}{x_{i0}}\right)}{dt} = u^i(R) - u^i(0)$ where $u^i(R), u^i(0)$ depend only on the mixed strategies of the rest of the agents (*i.e.*, other than i). As a result, the flow $\Psi^t = g \circ \Phi^t \circ g^{-1}$, which arises from our system via the change of variables $z_{iR} = \ln(x_{iR}/x_{i0})$, defines a separable vector field in the sense that the evolution of z_{iR}, depends only on the state variables of the other agents. The diagonal of the Jacobian of this vector field is zero and consequently the divergence (which corresponds to the trace of the Jacobian) is zero as well. Liouville's theorem states that such flows are volume preserving.

If we assume that replicator exhibits a uniformly persistent property $\Gamma \subset \mathcal{S}$ (for some $\epsilon > 0$), we clearly reach a contradiction. Indeed, consider the set of all (interior) points whose distance[7] from the boundary is at least $\epsilon/2$. Their corresponding limit sets will eventually converge to a smaller set whose distance from the boundary is at least ϵ. This is impossible due to conservation of volume of the conjugate flow. $\square$

We will identify necessary conditions for the existence of persistent properties in any graphical game and we will show how to decide them efficiently.

THEOREM 4. *Let Φ denote the flow of the replicator dynamic when applied to a graphical game and let the observation function $\mathcal{F}$ be the identity function. If $\times_i int(\Delta(S_i))$ is a persistent property of the flow then the flow has an interior fixed point q.*

PROOF. Let's pick an arbitrary initial condition $x_0 \in \text{int}(\times_i \Delta(S_i))$. If $\times_i \text{int}(\Delta(S_i))$ is a persistent property of the flow then by definition we have that there exists[8] $\epsilon_{x_0} > 0$ such that $\liminf \text{dist}(\Phi(x_0, t), \text{bd}(\times_i \Delta(S_i))) = \epsilon_{x_0}$, where for convenience here we take dist to denote the infinity norm. We denote as $x_i = \Phi_i(x_0, t)$ the vector encoding the mixed strategy of agent i over her available actions at time t. By assumption of persistence we have that there exists $\epsilon > 0$ and T_0 such that for all $t > T_0 : x_{iR}(t) > \epsilon$ for each agent i and strategy $R \in S_i$. We have that $\int_{T_0}^t \left[u^i(R) - \sum_{R \in S_i} x_{iR}u^i(R)\right]d\tau = \int_{T_0}^t \frac{\dot{x}_{iR}}{x_{iR}}d\tau = \ln\left(\frac{x_{iR}(t)}{x_{iR}(T_0)}\right)$. Furthermore, $\lim_{t\to\infty} \frac{1}{t} \ln\left(\frac{x_i(t)}{x_i(T_0)}\right) = 0$. For any pair of agent i and strategy R, the functions $\frac{1}{t}\int_{T_0}^t x_{iR}d\tau$, $\frac{1}{t}\int_{T_0}^t u^i(R)d\tau$ are bounded. Since they are finitely many of them we can find a common converging subsequence t_n for all of them[9]. Combining the last two equations and dividing them with t_n we derive for every agent i, $R \in S_i$:

$$\lim_{n\to\infty} \frac{1}{t_n} \int_{T_0}^t \sum_{R \in S_i} x_{iR}u^i(R)d\tau = \lim_{n\to\infty} \frac{1}{t_n} \int_{T_0}^t u^i(R)d\tau =$$

$$= \quad \lim_{n\to\infty} \frac{1}{t_n} \int_{T_0}^t \text{E}_{s_{-i}\sim x_{-i}(\tau)} u_i(R, s_{-i})d\tau = u_i(R, \hat{x}_{-i})$$

where $\hat{x}_{iR} = \lim_{n\to\infty} \frac{1}{t_n} \int_{T_0}^t x_{iR}d\tau$ and the last equation follows from the separability of payoffs. Since for all agents i, $\forall R, Q \in S_i : u_i(R, \hat{x}_{-i}) = u_i(Q, \hat{x}_{-i})$, $\hat{x}$ is a fully mixed Nash equilibrium. $\square$

Next, we will argue that this condition can be checked efficiently. That is, we will show that we can decide whether a graphical game has a fully mixed Nash equilibrium efficiently. Actually, we prove the following slightly stronger statement.

PROPOSITION 5. *Given any graphical game and a set of strategies $S'_i \subset S_i$ for each agent i we can decide whether it has a Nash equilibrium where each agent's i support is exactly S'_i.*

PROOF. For each agent i and strategy $s_i \in S'_i$ we solve the following related LP:

$$\begin{aligned}
&\max 1_{s_i} \cdot x_i \\
1_{i_1}^T \sum_{(i,k)\in E} A^{i,k}x_k = \quad \dots \quad &= 1_{i_{|S'_i|}}^T \sum_{(i,k)\in E} A^{i,k}x_k \;\forall i \\
1_{i_1}^T \sum_{(i,k)\in E} A^{i,k}x_k \quad &\geq \quad 1_{s_i}^T \sum_{(i,k)\in E} A^{i,k}x_k \;\forall i, \forall s_i \in S_i \\
&x_i \in \Delta(S'_i) \;\forall i
\end{aligned}$$

where 1_{i_k} is a vector of size $|S_i|$ where the i_k-th entry is equal to 1 and all the others are zero. Each such LP chooses from the set of all possible equilibria with the target supports one where the probability that agent i assigns to strategy $s_i \in S'_i$ is maximal. Naturally if the value of any of these LPs is equal to zero then there is no equilibrium with exactly the target support and we are done. If all of the values are positive then there exists an equilibrium with exactly the target support. This is because any convex combination of the solutions above is still a Nash equilibrium of our game. $\square$

Next, we will show that the existence of a fully mixed Nash equilibrium is also a sufficient condition for persistence in a large classes of graphical (polymatrix) games.

[5]Such techniques were first introduced by Hofbauer [10, 9].
[6]The reverse map is $x_{i0} = \frac{1}{1+\sum_{i\in S_i\setminus\{0\}} e^{z_{iR}}}$, $x_{iR} = \frac{e^{z_{iR}}}{1+\sum_{R\in S_i\setminus\{0\}} e^{z_{iR}}}$ for $R \in S_i \setminus \{0\}$. In fact, g is a diffeomorphism.
[7]Here, we use the infinity norm, however, this is not critical and we could use some other metric.
[8]This ϵ_{x_0} is a function of the initial condition x_0
[9]Take a convergent subsequence of the first function and find on this a convergent subsequence of the second and so on.

THEOREM 6. *Let Φ denote the flow of the replicator dynamic when applied to an affine transformation of zero-sum graphical game and let the observation function $\mathcal{F}$ be the identity function. If the flow has an interior fixed point q*[10] *then $\times_i int(\Delta(S_i))$ is a persistent property of the flow. Any property $\Gamma \subset \times_i\Delta(S_i)$ whose complement $\times_i\Delta(S_i)\backslash\Gamma$ contains a ball of radius $\epsilon > 0$ is not persistent.*

We will break down its proof into two lemmas. The first lemma aims to argue boundedness of the conjugate volume-preserving flow. The second lemma applies Poincaré recurrence theorem to argue the recurrent nature of the flow trajectories. This technique was introduced in [17] for proving the analogous statement in the case of separable zero-sum multiplayer games. In our case, we need to suitably update the two technical lemmas.

The first lemma identifies an information theoretic property that is invariant for the replicator. Similar invariants, without the connections to information theory though, have been established by Akin and Losert in special cases of zero-sum games [1].

LEMMA 7. *Let Φ denote the flow of the replicator dynamic when applied to a $(\vec{a},\vec{b})$-zero-sum multiplayer game G and let GG be the separable zero-sum game defined by the utility functions $u_i'(s) = a_iu_i + b_i$. If GG (or equivalently G) has an interior (i.e. fully mixed) Nash equilibrium $q = (q_1, q_2, \ldots, q_n)$ then given any starting point $x_0 \in \times_i int(\Delta(S_i))$, $\sum_i a_iH(q_i, \Phi_i(x_0, 0)) = \sum_i a_iH(q_i, \Phi_i(x_0, t))$ for all $t \geq 0$.*

PROOF. We will show that the derivative of the quantity $\sum_i a_iH(q_i, \Phi_i(x_0, t)) = -\sum_i a_i \sum_{R\in S_i} q_{iR}\cdot\ln(x_{iR})$ is everywhere zero.

$$\begin{aligned}
\sum_i a_i \sum_{R\in S_i} q_{iR}\frac{d\ln(x_{iR})}{dt} &= \sum_i a_i \sum_{R\in S_i} q_{iR}\frac{\dot{x}_{iR}}{x_{iR}} = \\
&= \sum_i a_i\big(\sum_{R\in S_i} q_{iR}u^i(R) - \sum_{R\in S_i} x_{iR}u^i(R)\big) = \\
&= \sum_i \big(\sum_{R\in S_i} q_{iR}(a_iu^i(R)+b_i) - \sum_{R\in S_i} x_{iR}(a_iu^i(R)+b_i)\big) = \\
&= \sum_i \sum_{(i,k)\in E} \big(q_i^{\mathrm{T}}A^{i,k}x_k - x_i^{\mathrm{T}}A^{i,k}x_k\big) = \\
&= \sum_i \sum_{(i,k)\in E} \big(q_i^{\mathrm{T}} - x_i^{\mathrm{T}}\big)A^{i,k}(x_k - q_k) = \\
&= -\sum_{(i,k)\in E, i<k} \big[\big(q_i^{\mathrm{T}} - x_i^{\mathrm{T}}\big)A^{i,k}(q_k - x_k) + \\
&\qquad + \big(q_i^{\mathrm{T}} - x_i^{\mathrm{T}}\big)\big(c_{\{i,k\}}\mathbf{1} - A^{i,k}\big)(q_k - x_k)\big] = 0
\end{aligned}$$

We have used the fact that for each agent i, $\sum_{(i,k)\in E}\big(q_i^{\mathrm{T}} - x_i^{\mathrm{T}}\big)A^{i,k}q_k = u_i^{GG}(q) - u_i^{GG}(x_i, q_{-i}) = 0$ since q is a fully mixed Nash. $\square$

COROLLARY 8. *The weighted sum of the Kullback-Leibler divergences of each agent's i current strategy from q_i is a constant of the motion for the flow. Equivalently, given any starting point $x_0 \in \times_i int(\Delta(S_i))$, $\sum_i a_iD_{\mathrm{KL}}(q_i\|\Phi_i(x_0, 0)) = \sum_i a_iD_{\mathrm{KL}}(q_i\|\Phi_i(x_0, t))$ for all $t \geq 0$.*

We simplify notation by using $D^a_{\mathrm{KL}}(q\|x)$ in the place of $\sum_i a_iD_{\mathrm{KL}}(q_i\|x_i)$. We have established that the replicator flow is conjugate to a volume preserving flow from the proof of theorem 3. On the other hand, the applied transformation blows up the volume near the boundary to infinity and as a result does not allow for an immediate application of Poincaré's recurrence theorem. We circumvent this issue by applying corollary 8.

LEMMA 9. *If the flow Φ has an interior fixed point then given any open set E that is bounded away from $bd(\times_i\Delta(S_i))$ there exist orbits that intersect it infinitely often.*

PROOF. Let q be the interior fixed point of the flow. Given any open set E that is bounded away from the boundary and let $c_E = \sup_{x\in E} D^a_{\mathrm{KL}}(q\|x)$. Since E is bounded away from the boundary c_E is finite. We focus on the restriction of conjugate flow Ψ over the closed and bounded set[11] $g(\mathcal{S}_E)$, where $\mathcal{S}_E = \{x \in \times_i\Delta(S_i) : D^a_{\mathrm{KL}}(q\|x) \leq c_E\}$. The fact that replicator preserves weighted K-L divergences implies that replicator maps $\mathcal{S}_E$ to itself. Due to the homeomorphism g, the same applies for flow Ψ and $g(\mathcal{S}_E)$. The restriction of flow Ψ on $g(\mathcal{S}_E)$ is a volume preserving flow and has only bounded orbits. We can apply Poincaré's recurrence theorem to derive that for each open set of this system there exist orbits $\Psi(z_0, \cdot)$ that intersect this set infinitely often. Given our initial arbitrary (but bounded from the boundary of $\times_i\Delta(S_i))$) open set E, $g(E)$ is also open[12] and hence infinitely recurrent for some $\Psi(z_0, \cdot)$ but now the $g^{-1}(\Psi(z_0, \cdot)) = \Phi(g^{-1}(z_0), \cdot)$ visits E infinitely often, concluding the proof. $\square$

We will combine lemmas 7 and 9 to prove theorem 6.

PROOF. Since, the weighted K-L divergence between the state of the system and the fully mixed Nash equilibrium remains constant we have that for each initial condition x_0 the trajectory $\Phi(x_0, t)$ must stay bounded away from the boundary since there the weighted K-L divergence becomes infinite. This implies that for each $x_0 \in \times_i\mathrm{int}(\Delta(S_i))$ there exists[13] $\epsilon > 0$ such that $\liminf \mathrm{dist}(\Phi(x_0, t), \mathrm{bd}(\times_i\Delta(S_i))) = \epsilon > 0$, therefore $\times_i\mathrm{int}(\Delta(S_i))$ is a persistent property of the flow. For any (nontrivial) property $\Gamma \subset \times_i\Delta(S_i)$ whose complement contains a ball of radius $\epsilon > 0$, by lemma 9, we have established that there exist initial conditions whose trajectories revisit these points infinitely often and therefore Γ is not persistent. $\square$

We have already shown that we can check whether any graphical polymatrix games has a fully mixed equilibrium efficiently. Another interesting question is whether we can check efficiently whether a graphical polymatrix games corresponds to an (affine) variant of separable multi-agent zero-sum game. This is an interesting question since the set of all possible strategy outcomes has exponential size and therefore we cannot inspect every single outcome individually.

LEMMA 10. *Given any graphical polymatrix game G and a set of affine transformations of utilities, we can check efficiently whether this pair corresponds to an affine transformation of a separable zero-sum multi-agent game.*

[10] This is equivalent to game having fully mixed Nash equilibrium.

[11] $g(\mathcal{S}_E)$ is closed since $\mathcal{S}_E$ is closed and g is a homeomorphism. $g(\mathcal{S}_E)$ is bounded since E is bounded away from the boundary of $\times_i\Delta(S_i)$.

[12] Since g is a homeomorphism.

[13] This ϵ is a function of the initial condition x_0.

PROOF. The proof of the lemma relies on the observation of [7], where the authors show that given any separable zero-sum multi-agent game, there exists a poly-time computable payoff preserving transformation of this game to a graphical polymatrix game where each edge/game is a constant-sum game (with possibly different constants on each edge). Our argument works in two steps, first we apply the reverse affine transformation to G so as to transform it to a separable zero-sum game. Next, we apply the payoff preserving transformation to produce it in its network constant-sum form. Finally, this property is trivially testable in polynomial size, by verifying that each edge game is indeed constant-sum. □

5.2 Weak Persistence

In this section, we wish to contrast the sparsity of (strongly) persistent properties in graphical polymatrix games (for the replicator flow) against the abundance of weakly persistent properties. It suffices to focus on the simplest possible graphical polymatrix games. Indeed, we will show that even the classic game of Matching Pennies has uncountably infinite weakly persistent properties under the replicator dynamic. Furthermore, the intersection of any two such properties can be arbitrarily small. Since Matching Pennies is a zero-sum game and has a (unique) fully mixed equilibrium its only (interior) persistent property is the whole interior of the state space. Such statements would have been impossible under the assumption of convergence to equilibrium. Nevertheless, they are useful and meaningful. They encode a system that act as a reservoir of an infinite collection of distinct information theoretic properties.

THEOREM 11. *Let Φ denote the flow of the replicator dynamic when applied to Matching Pennies game and let the observation function $\mathcal{F}$ be the projection of the state space to the subspace encoding the probabilities that each agent assigns to action "Head". Any open set $\Gamma \subset (0,1)^2$ containing $S_\alpha = \{(x_1, x_2) : x_2 = \alpha(x - \frac{1}{2}) + 1/2, 0 < x_1 < 1, 0 < x_2 < 1\}$ for some $\alpha \in \mathbb{R}$ is weakly persistent.*

PROOF. We will show that given any such property $\Gamma \subset (0,1)^2$ for all initial conditions $x \in \text{int}(\mathcal{S})$ we have that $\liminf_{t\to\infty} \text{dist}(\mathcal{F}(\Phi(x,t)), O\backslash\Gamma) > 0$. It suffices to show that for all initial conditions $x \in \text{int}(\mathcal{S})$ each set S_α is infinitely recurrent, *i.e.*, $\sup\{t : \Phi(x,t) \in S_\alpha\} = \infty$. Suppose that this was not the case. Let $S_\alpha^+ = \{(x_1, x_2) : x_2 > \alpha(x - \frac{1}{2}) + 1/2, 0 < x_1 < 1, 0 < x_2 < 1\}$ and $S_\alpha^- = \{(x_1, x_2) : x_2 < \alpha(x - \frac{1}{2}) + 1/2, 0 < x_1 < 1, 0 < x_2 < 1\}$. If the set S_α is not infinitely recurrent, then there exists a trajectory whose limit points exclusively in either S_α^+ or S_α^-. However, as we have seen in the proof of theorem 4 the time averages of the agent strategies have a convergent subsequent whose limit is the unique Nash. The only way that this would be possible was if the trajectory converged to the equilibrium, however, this is impossible due to the KL-divergence invariance. □

Of course, this is not the unique (infinite) class of persistent properties. For example, instead of using lines to define partitions of the state space, we could use other families of smooth curves that satisfy the unique Nash equilibrium and essentially the same proof would carry over.

We conclude the analysis by excluding the possibility of uniformly weakly persistent properties. for Matching Pennies games $\Gamma \subset (0,1)^2$ when the observation observation function $\mathcal{F}$ is the projection of the state space to the subspace encoding the probabilities that each agent assigns to

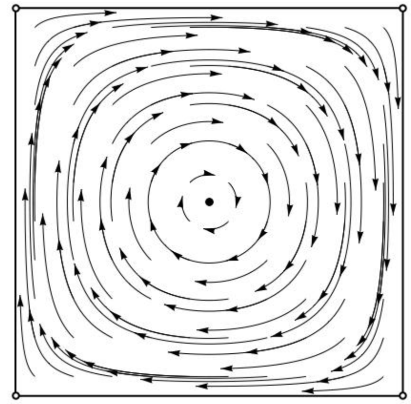

Figure 2: Replicator orbits in Matching Pennies (probabilities of each agent playing heads)

action "Head". The complement of any such property contains the boundary of $[0,1]^2$, however, as we will argue in the following section for any $\epsilon > 0$ there exist periodic orbits such that each of their points is at a distance less than ϵ from the boundary (see figure 2). This excludes the possibility of uniformly weakly persistent properties in $(0,1)^2$.

COROLLARY 12. *Let Φ denote the flow of the replicator dynamic when applied to the Matching Pennies game and let the feature $\mathcal{F}$ be the projection of the state space to the subspace encoding the probabilities that each agent assigns to action "Head". The flow has no uniformly weakly persistent property $\Gamma \subset (0,1)^2$ for $\mathcal{F}$.*

5.3 Patterns

We close our discussion by establishing some detailed properties of the replicator flow of a dynamic nature.

THEOREM 13. *Let Φ denote the flow of the replicator dynamic when applied to the Matching Pennies game and let the observation function $\mathcal{F}$ be the projection of the state space to the subspace encoding the probabilities that each agent assigns to action "Head". Every starting point (x_0, y_0), other than the equilibrium, lies on a periodic orbit. The equation of the limit cycle is $\{(x^{Head}, y^{Head}) : x^{Head}(1 - x^{Head})y^{Head}(1 - y^{Head}) = x_0^{Head}(1 - x_0^{Head})y_0^{Head}(1 - y_0^{Head}), 0 < x^{Head}, y^{Head} < 1\}$ and its direction is clockwise.*

PROOF. We know that in zero-sum games with fully mixed equilibria the KL-divergence between the Nash equilibrium and the state of the system remains constant as we move along the trajectories of the replicator. KL-divergence is a (pseudo)-metric implying the existence of trapping regions in the interior of $[0,1]^2$. Specifically, as long as we start from an interior point other than the unique Nash then the trajectory stays bounded away from the boundary (KL-divergence becomes infinite) and from the unique equilibrium (KL-divergence becomes zero). By the Poincaré-Bendixson theorem we have that starting from any point (other than the Nash) the resulting limit set is a periodic orbit. It is straightforward to check that the KL-divergence invariance condition when projected to our subspace translates to an invariance of the quantity $x^{Head}(1-x^{Head})y^{Head}(1-y^{Head})$. This defines a closed continuous curve on our subspace that is symmetric along the axis $x^{Head} = 1/2$ and $y^{Head} = 1/2$. In each of the four regions defined by $x^{Head} = 1/2$, $y^{Head} = 1/2$ the signs of $\frac{dx^{Head}}{dt}$, $\frac{dy^{Head}}{dt}$ are fixed and define a clockwise direction. The theorem follows from the uniqueness of the solution of the replicator flow. $\square$

6. DISCUSSION

What is a persistent (and efficient) reservoir of useful properties/patterns? *A (well-designed) multi-agent system.* A multi-agent system is expressed via the collection of persistent spatiotemporal correlations amongst its numerous, dispersed members. These correlations can be of a static (equilibria) or a dynamic nature. Regardless of their specifics it is the pursuit of these long range correlations that necessitates the employment of multi-agent learning.

A solution may additionally exhibit other (generic/context specific) desirable properties. Examples of generic properties could be low computational, randomness, or communication costs on the side of the agents, whereas context-specific could be high social welfare in the case of socioeconomic systems, or high probability of safe long term operation for a team of deployed robots.

Persistent pattern implementation, via multi-agent system design, arises as a largely unexplored area that holds the promise of renewing our approach to multi-agent systems. However, which patterns should we attempt to replicate? A bottom-up approach, where we start with some simple parametric family of patterns that can be weaved together to produce more complicated ones, seems promising.

7. ACKNOWLEDGEMENTS

This work was generously supported by grants AFOSR MURI (FA9550-09-1-0538), ONR project (N00014-09-1-0751) and the Army Research Lab (ARL) MAST CTA project (329420). The authors would also like to thank our reviewers for their helpful comments and suggestions.

8. REFERENCES

[1] E. Akin and V. Losert. Evolutionary dynamics of zero-sum games. *J. of Math. Biology*, 20:231–258, 1984.

[2] B. Alpern and F. Schneider. Recognizing safety and liveness. *Distributed Computing*, 2(3):117–126, 1987.

[3] L. Barreira. Poincare recurrence: old and new. In *XIVth International Congress on Mathematical Physics. World Scientific.*, pages 415–422, 2006.

[4] I. Bendixson. Sur les courbes définies par des équations différentielles. *Acta Mathematica*, 24(1):1–88, 1901.

[5] N. P. Bhatia and G. P. Szegö. *Stability theory of dynamical systems.* Berliin: Springer Verlag, 1970.

[6] G. Brown. Iterative solutions of games by fictitious play. *In Activity Analysis of Production and Allocation, T.C. Koopmans (Ed.), New York: Wiley.*, 1951.

[7] Y. Cai and C. Daskalakis. On minmax theorems for multiplayer games. In *ACM-SIAM Symposium on Discrete Algorithms*, SODA, pages 217–234, 2011.

[8] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* Wiley, New York, 1991.

[9] J. Hofbauer. Evolutionary dynamics for bimatrix games: A hamiltonian system? *J. of Math. Biology*, 34:675–688, 1996.

[10] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics.* Cambridge University Press, Cambridge, 1998.

[11] V. Hutson and S. K. Permanence and the dynamics of biological systems. *Math. Biosciences*, 111:293–326, 1992.

[12] R. Kleinberg, K. Ligett, G. Piliouras, and É. Tardos. Beyond the Nash equilibrium barrier. In *Symposium on Innovations in Computer Science (ICS)*, 2011.

[13] R. Kleinberg, G. Piliouras, and É. Tardos. Multiplicative updates outperform generic no-regret learning in congestion games. In *ACM Symposium on Theory of Computing (STOC)*, 2009.

[14] L. Lamport. Proving the correctness of multiprocess programs. *Software Engineering, IEEE Transactions on*, SE-3(2):125–143, March 1977.

[15] S. Mannor and J. S. Shamma. Multi-agent learning for engineers. *Artificial Intelligence*, 171(7):417 – 422, 2007.

[16] J. Meiss. *Differential Dynamical Systems.* SIAM, 2007.

[17] G. Piliouras and J. S. Shamma. Optimization despite chaos: Convex relaxations to complex limit sets via Poincaré recurrence. In *Symposium of Discrete Algorithms (SODA)*, 2014.

[18] H. Poincaré. Sur le problème des trois corps et les équations de la dynamique. *Acta Math*, 13:1■270, 1890.

[19] J. Robinson. An iterative method of solving a game. *Annals of Mathematics*, 54:296–301, 1951.

[20] W. H. Sandholm. *Population Games and Evolutionary Dynamics.* MIT Press, 2010.

[21] W. H. Sandholm, E. Dokumaci, and F. Franchetti. Dynamo: Diagrams for evolutionary game dynamics, 2012.

[22] P. Schuster and K. Sigmund. Replicator dynamics. *Journal of Theoretical Biology*, 100(3):533 – 538, 1983.

[23] Y. Shoham, R. Powers, and T. Grenager. If multi-agent learning is the answer, what is the question? *Artificial Intelligence*, 171(7):365–377, 2007.

[24] P. Stone. Multiagent learning is not the answer. it is the question. *Artificial Intelligence*, 171(7):402 – 405, 2007.

[25] P. D. Taylor and L. B. Jonker. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences*, 40(1âĂŞ2):145 – 156, 1978.

[26] G. Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Soc., 2012.

[27] R. V. Vohra and M. P. Wellman. Foundations of multi-agent learning: Introduction to the special issue. *Artificial Intelligence*, 171(7):363 – 364, 2007.

[28] J. von Neumann. Zur theorie der gesellschaftsspiele. *Mathematische Annalen*, 100:295–300, 1928.

[29] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, 1944.

[30] J. W. Weibull. *Evolutionary Game Theory.* MIT Press; Cambridge, MA: Cambridge University Press., 1995.