Robust Self-organization in Games: Symmetries, Conservation Laws and Dimensionality Reduction

Extended Abstract

Sai Ganesh Nagarajan
SUTD
Singapore
sai_nagarajan@mymail.sutd.edu.sg

David Balduzzi
Google DeepMind
London, UK
dbalduzzi@google.com

Georgios Piliouras
SUTD
Singapore
georgios@sutd.edu.sg

ABSTRACT

Games are an increasingly useful tool for training and testing learning algorithms. Recent examples include GANs, AlphaZero and the AlphaStar league. However, multi-agent learning can be extremely difficult to predict and control. Learning dynamics can yield chaotic behavior even in simple games. In this paper, we present basic *mechanism design* tools for constructing games with predictable and controllable dynamics. We present a robust framework for dimensionality reduction arguments in large network games.

KEYWORDS

[COIN] Self-organisation; [LAD] Multi-agent learning; [MANG] Non-cooperative games: theory & analysis

ACM Reference Format:

Sai Ganesh Nagarajan, David Balduzzi, and Georgios Piliouras. 2020. Robust Self-organization in Games: Symmetries, Conservation Laws and Dimensionality Reduction. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Games have become a powerful training mechanism used to generate photorealistic images [4], and also to play Go, Chess and StarCraft [1, 6, 13–15]. The underlying assumption behind this system architecture is that competition between learning algorithms forces them to continually improve their performance. However, this assumption is not always valid in practice. Indeed, in numerous cases of multi-agent competition (or even cooperation) the resulting dynamics can be unpredictable or, even worse, formally chaotic [3, 10–12]. This raises our central problem: *How can we effectively control learning dynamics in games?*

In this paper, we present a new perspective on mechanism design, that is concerned with the *dynamics of learning in games* rather than fixed-points/equilibria.

Our approach. We define a new approach to solving games. More precisely we design a novel framework for articulating when a multi-agent learning system has succeeded in becoming coordinated. We do this by circumventing equilibration, i.e., we do not artificially require our system to fixate. Instead, we examine conditions under which non-equilibrating dynamics in games can exhibit regularities such as conservation laws (invariant functions) as well as periodicity.

Structure. In section 2 we describe our exact setting, which includes a new class of polymatrix [7]) n-player games, that we call **network constant-sum games with charges**. We assume that agents apply (possibly different variants) of *Follow-the-Regularized-Leader (FTRL) dynamics [5, 9].* In section 3, we prove conservation laws for the dynamics and show how symmetries in these games can be used for dimensionality reduction. In section 4, we prove periodicity for a specific subclass of network constant-sum games with charges. Section 5 shows how given a target conservation law one can reverse-engineer a game that implements it.

2 PRELIMINARIES

2.1 Network Zero-Sum Games with Charges

A graphical polymatrix game is defined by an undirected graph $G = (V, E)$, where V corresponds to the set of agents and where edges correspond to bimatrix games between the 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 then we denote by $s_i \in S_i$ the respective strategy of agent i. Similarly, let $s_{-i} \in \times_{j \in V \setminus i} S_j$ denote the strategies of the other agents. The payoff of agent $i \in V$ in strategy profile s is the sum of the payoffs that agent i receives from all the bimatrix games she participates in. Specifically, $u_i(s) = \sum_{(i,k)\in E} A^{i,k}_{s_i,s_k}$. A randomized strategy x_i for agent i lies on the simplex $\Delta(S_i) = \{p \in \mathbb{R}_+^{|S_i|} : \sum_{R\in S_i} x_{iR} = 1\}$. Payoff functions are extended to randomized strategies in the usual multilinear fashion. A (mixed) Nash equilibrium is a profile of mixed strategies such that no agent can improve her payoff by unilaterally deviating to another strategy.

Definition 2.1. [2] A separable constant-sum multiplayer game GG is a graphical polymatrix game in which, for any pure strategy profile, the sum of all agent payoffs is equal to the same constant. Formally, $\forall s \in \times_i S_i, \sum_i u_i(s) = c$.

Our main class of games will be produced by taking linear transformations (rescalings with possible switch of the direction of axes) of separable zero-sum games of the form $\lambda_i u_i$, where $\lambda_i \in \mathbb{R} \setminus \{0\}$. That is we can think of each agent as a charged particle, where their charge, λ_i can be either positive or negative.

Definition 2.2. An n-agent game G is a network constant-sum game with charges if there exists a separable constant-sum multiplayer game GG and constants $\lambda_i \in \mathbb{R} \setminus \{0\}$ for each agent i such

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

that $u^{GG}(s) = \lambda_i u_i^G(s)$ for each outcome $s \in S$. We will also denote such game as $(\vec{\lambda})$-constant-sum multiplayer game.

2.2 Follow the Regularized Leader

Follow the Regularized Leader (FTRL) is a class of learning dynamics that tries to optimize the strategies being played by tracking cumulative payoffs over time and trying to maximize the payoff to each player at every time instant. To this end, let $v_{iR}(x) := u_i(R, x_{-i})$ and thus $v_i(x) = (v_{iR}(x))_{R \in S_i}$ The FTRL dynamics can be specified as follows:

$$y_i(t) = y_i(0) + \int_0^t v_i(x(s))ds, \tag{FTRL}$$

$$x_i(t) = Q_i(y_i(t)), \tag{1}$$

where $Q_i : \mathbb{R}^{S_i} \mapsto \mathcal{X}_i$ is defined as

$$Q_i(y_i) = \arg\max_{x_i \in \mathcal{X}_i} \{\langle y_i, x_i \rangle - h_i(x_i)\}. \tag{2}$$

Here $\mathcal{X}_i = \Delta(S_i)$.

Another useful notion is of the convex conjugate of $h_i(x)$, which is defined to be:

$$h_i^*(y_i) = \max_{x_i \in \mathcal{X}_i} \{\langle y_i, x_i \rangle - h_i(x_i)\}. \tag{3}$$

3 DIMENSIONALITY REDUCTION AND INVARIANT FUNCTIONS

3.1 Constant of Motion

We describe a function that is invariant to the evolution of the agents' strategies over time, when playing a network constant-sum game with charges, i.e., a constant of motion. We show that the time derivative of $H(y) := \sum_{i \in V} \lambda_i \left(h_i^*(y_i) - \langle y_i, x_i^* \rangle\right)$ is zero, i.e, $H(y)$ remains invariant with the motion of the FTRL dynamics. In the above definition x^* is an interior Nash equilibrium.

Theorem 3.1. *$H(y) := \sum_{i \in V} \lambda_i \left(h_i^*(y_i) - \langle y_i, x_i^* \rangle\right)$ is invariant to the evolution of FTRL dynamics when agents play a network constant-sum game with charges.*

3.2 Dimensionality Reduction & Symmetries

Consider a network constant-sum game with charges where each edge game is the same n-by-n zero-sum (base) game with payoff matrix A. The network topology is a bipartite graph with L layers and each layer has K vertices (agents). Edges/games exist only between successive layers and the corresponding subgraph is complete (i.e., all agents in layer j play against all agents in layer $j+1$). Let the set of all vertices be V and the set of all edges be E. The agents are indexed by their vertex and the layer. For instance the agent in vertex i and layer j is indexed as (i, j). The corresponding mixed strategies, payoff vectors and utilities are thus going to be indexed by $x_{(i,j)}$, $y_{(i,j)}$, $u_{(i,j)}$ and the vector of charges by $\vec{\lambda} = \left[\lambda_{(1,1)}, \lambda_{(2,1)}, \ldots, \lambda_{(L,K)}\right]$.

We use $\mathcal{G} := \left(V, E; \vec{\lambda}, A\right)$ to represent this setting that is parameterized by charges $\vec{\lambda}$ and the base game matrix A. Then the following theorem holds:

Theorem 3.2. *The dynamical system induced by agents using any FTRL dynamics in any symmetric bipartite network constant-sum game with charges $\mathcal{G} = \left(V, E; \vec{\lambda}, A\right)$, lies on a low dimensional space requiring only $L(n-1)$ variables to completely describe the system.*

4 PERIODIC ORBITS

Here we show how we can leverage the dimensionality reduction proven in the previous section to establish the emergence of periodic orbits and other useful properties about the system dynamics (such as the lack of chaos).

Theorem 4.1. *If the setting of symmetric bipartite network constant-sum games with charges $\mathcal{G} = \left(V, E; \vec{\lambda}, A\right)$ consists of two layers and each agent has two actions and the charges of all agents have the same sign then almost all orbits are periodic.*

Thus, we can understand to a large extent the topology of these multi-agent systems, despite the fact that they correspond to games with possibly arbitrarily large number of agents. When contrasting this with the possibility of chaos even in two player games [10, 12], we see the power of these techniques.

5 REVERSE-ENGINEERING THE GAME

Our stated goal is to design systems that exhibit conservation laws. By theorem 3.1 we know we can enforce a parametric family of constants of motions of the form $H(y) := \sum_{i \in V} \lambda_i \left(h_i^*(y_i) - \langle y_i, x_i^* \rangle\right)$. Here we explore the reverse question. Can we efficiently find a game that implements such a law?

Theorem 5.1. *Given any conservation law of the form $H(y) := \sum_{i \in V} \lambda_i \left(h_i^*(y_i) - \langle y_i, x_i^* \rangle\right)$ we can compute in linear time a network constant-sum game with charges that implements it where each agent i uses FTRL dynamics with regularizer h_i. Moreover, the payoff matrices of the network constant-sum game with charges are sparse.*

6 CONCLUSION

What is self-organization? We know it when we see it in familiar games like soccer, where forcing teams to compete encourages players to learn coordinated behaviors such as passing [8]. In this paper, we precisely characterize how self-organization arises in network constant-sum games with charges. Our strategy is twofold. Firstly, we show that the games satisfy conservation laws, i.e, the dynamics of the game are contained in level sets of certain invariant functions and hence live on a (sometimes much) lower dimensional subspace of the space of possible joint actions. Secondly, we apply the dimensionality reduction argument to show that, for symmetric games on a bipartite graph, the limit behaviors of the dynamics are simple (periodic) and chaotic dynamics are excluded.

ACKNOWLEDGMENTS

Sai Ganesh Nagarajan would like to acknowledge the SUTD President's Graduate fellowship. Georgios Piliouras acknowledges MOE AcRF Tier 2 Grant 2016-T2-1-170, grant PIE-SGP-AI-2018-01 and NRF 2018 Fellowship NRF-NRFF2018-07.

REFERENCES

[1] Trapit Bansal, Jakub Pachocki, Szymon Sidor, Ilya Sutskever, and Igor Mordatch. 2018. Emergent complexity via multi-agent competition. *ICLR* (2018).

[2] Yang Cai and Costantinos Daskalakis. 2011. On Minmax Theorems for Multiplayer Games. In *ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 217–234.

[3] Tobias Galla and J Doyne Farmer. 2013. Complex dynamics in learning complicated games. *Proceedings of the National Academy of Sciences* 110, 4 (2013), 1232–1236.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In *Advances in neural information processing systems*. 2672–2680.

[5] Elad Hazan et al. 2016. Introduction to online convex optimization. *Foundations and Trends® in Optimization* 2, 3-4 (2016), 157–325.

[6] Max Jaderberg, Wojciech M Czarnecki, Iain Dunning, Luke Marris, Guy Lever, Antonio Garcia Castaneda, Charles Beattie, Neil C Rabinowitz, Ari S Morcos, Avraham Ruderman, Nicolas Sonnerat, Tim Green, Louise Deason, Joel Z. Leibo, David Silver, Demis Hassabis, Koray Kavukcuoglu, and Thore Graepel. 2018. Human-level performance in first-person multiplayer games with population-based deep reinforcement learning. *arXiv:1807.01281* (2018).

[7] M. J. Kearns, M. L. Littman, and S. P. Singh. 2001. Graphical Models for Game Theory. In *UAI*.

[8] Siqi Liu, Guy Lever, Josh Merel, Saran Tunyasuvunakool, Nicolas Heess, and Thore Graepel. 2019. Emergent Coordination Through Competition. In *ICLR*.

[9] Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. 2018. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, 2703–2717.

[10] Gerasimos Palaiopanos, Ioannis Panageas, and Georgios Piliouras. 2017. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. In *Advances in Neural Information Processing Systems*. 5872–5882.

[11] G. Piliouras and J. S. Shamma. 2014. Optimization Despite Chaos: Convex Relaxations to Complex Limit Sets via Poincaré Recurrence. In *SODA*.

[12] Yuzuru Sato, Eizo Akiyama, and J. Doyne Farmer. 2002. Chaos in learning a simple two-person game. *Proceedings of the National Academy of Sciences* 99, 7 (2002), 4748–4751. https://doi.org/10.1073/pnas.032086299 arXiv:http://www.pnas.org/content/99/7/4748.full.pdf+html

[13] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. 2017. Mastering the game of Go without human knowledge. *Nature* 550 (2017), 354–359.

[14] Gerald Tesauro. 1995. Temporal difference learning and TD-Gammon. *Commun. ACM* 38, 3 (1995), 58–68.

[15] Oriol Vinyals, Igor Babuschkin, Junyoung Chung, Michael Mathieu, Max Jaderberg, Wojciech M. Czarnecki, Andrew Dudzik, Aja Huang, Petko Georgiev, Richard Powell, Timo Ewalds, Dan Horgan, Manuel Kroiss, Ivo Danihelka, John Agapiou, Junhyuk Oh, Valentin Dalibard, David Choi, Laurent Sifre, Yury Sulsky, Sasha Vezhnevets, James Molloy, Trevor Cai, David Budden, Tom Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Toby Pohlen, Yuhuai Wu, Dani Yogatama, Julia Cohen, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Chris Apps, Koray Kavukcuoglu, Demis Hassabis, and David Silver. 2019. AlphaStar: Mastering the Real-Time Strategy Game StarCraft II. https://deepmind.com/blog/alphastar-mastering-real-time-strategy-game-starcraft-ii/. (2019).