Learning Simulation-Based Games from Data

Extended Abstract

Enrique Areyan Viqueira
Brown University
Providence, Rhode Island, USA
NEC-AIST AI Cooperative Research Laboratory
Tokyo, Japan
eareyan@brown.edu

Amy Greenwald
Brown University
Providence, Rhode Island, USA
NEC-AIST AI Cooperative Research Laboratory
Tokyo, Japan
amy_greenwald@brown.edu

Cyrus Cousins
Brown University
Providence, Rhode Island, USA
cyrus_cousins@brown.edu

Eli Upfal
Brown University
Providence, Rhode Island, USA
eli_upfal@brown.edu

ABSTRACT

We tackle a fundamental problem in empirical game-theoretic analysis (EGTA), that of learning equilibria of simulation-based games. Such games cannot be described in analytical form; instead, a black-box simulator can be queried to obtain noisy samples of utilities. Our approach to EGTA is in the spirit of probably approximately correct learning. We design algorithms that learn empirical games, which uniformly approximate the utilities of simulation-based games from finitely many samples. Our methodology learns *all* the equilibria of simulation-based games, as opposed to a single one.

KEYWORDS

Empirical Game-Theoretical Analysis; PAC Learning

ACM Reference Format:
Enrique Areyan Viqueira, Amy Greenwald, Cyrus Cousins, and Eli Upfal. 2019. Learning Simulation-Based Games from Data. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

1 OVERVIEW

This paper is concerned with analyzing games for which a complete and accurate description is not available. While knowledge of the number of agents and their strategy sets is available, we do not assume *a priori* access to the game's utility functions. Instead, we assume access to a simulator from which we can sample noisy utilities associated with any strategy profile. Such games have been called *simulation-based games* [14] and *black-box games* [6], and their analysis is called *empirical game theoretic analysis* (EGTA) [2, 15]. EGTA methodology has been applied in various practical settings, including trading agent analyses in supply chain management [2, 13], ad auctions [4], and energy markets [5]; designing network routing protocols [16]; strategy selection in real-time games [9]; and the dynamics of reinforcement learning algorithms, like AlphaGo [10].

The aim of this work is to design learning algorithms that can accurately estimate *all* the equilibria of simulation-based games. We tackle this problem using the probably approximately correct (PAC) learning framework [11]. Our algorithms learn so-called *empirical games* [15], which are estimates of simulation-based games constructed via sampling. We prove that empirical games so constructed are *uniform* approximations of simulation-based games, meaning all utilities in the empirical game tend toward their expected counterparts in the simulation-based game, *simultaneously.*

This notion of uniform approximation is central to our work. Our main theorem states: when one game Γ is a uniform approximation of another Γ', all equilibria in Γ are approximate equilibria in Γ'. Intuitively, this theorem establishes perfect recall by the approximate game, in the sense that the approximate game contains all true positives: i.e., all (exact) equilibria of the original game. It also establishes approximately perfect precision, in the sense that all false positives in the approximate game are approximate equilibria in the original game. Our learning algorithms, which learn empirical games that are uniform approximations of simulation-based games, thus well estimate the equilibria of simulation-based games.

Related Work. One distinguishing feature of our work *vis à vis* the literature is that we aim to estimate *all* the equilibria of simulation-based games, rather than just a single one (e.g., [3]). Two notable exceptions are in the work by Vorobeychik [12] and Tuyls et al. [10]. The former includes asymptotic results about the equilibria of empirical games; we improve their analysis, showing finite-sample bounds. The latter derives guarantees on the quality of all equilibria learned from finite samples; beyond likewise establishing perfect recall, we further establish approximately perfect precision.

2 APPROXIMATING GAMES

We begin by presenting standard game-theoretic notions. We then define uniform approximation, and show that finding the approximate equilibria of a uniform approximation of a game is sufficient for finding all the (exact) equilibria of the game itself.

Definition 2.1 (Pure Normal-Form Game). A **pure normal-form game (NFG)** $\Gamma \doteq \langle \boldsymbol{P}, \{S_p\}_{p \in \boldsymbol{P}}, \boldsymbol{u}(\cdot)\rangle$ consists of a set of agents $\boldsymbol{P}$, with **pure strategy set** S_p for agent $p \in \boldsymbol{P}$. We define $\boldsymbol{S} \doteq S_1 \times \cdots \times S_{|\boldsymbol{P}|}$ to be the pure strategy profile space of Γ, and then $\boldsymbol{u} : \boldsymbol{S} \to \mathbb{R}^{|\boldsymbol{P}|}$ is a vector-valued utility function (equivalently, a vector of $|\boldsymbol{P}|$ scalar utility functions $\boldsymbol{u}_p$).

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

Given a pure NFG Γ, we denote by $S_p^\diamond$ the set of distributions over S_p; this set is called agent p's **mixed strategy set**. We define $\boldsymbol{S}^\diamond = S_1^\diamond \times \cdots \times S_{|\boldsymbol{P}|}^\diamond$, and then, overloading notation, we write $\boldsymbol{u}(s)$ to denote the expected utility of a mixed strategy profile $\boldsymbol{s} \in \boldsymbol{S}^\diamond$.

Definition 2.2 (ϵ-Nash Equilibrium). A pure (or mixed) strategy profile $\boldsymbol{s}$ in game Γ with utility function $\boldsymbol{u}$ is an ϵ-**Nash equilibrium (NE)** if $\sup_{p \in \boldsymbol{P}} \sup_{s' : s'_j = s_j \forall j \neq p} \boldsymbol{u}_p(s') - \boldsymbol{u}_p(s) \leq \epsilon$.

Given a game Γ, we denote by $\mathrm{E}_\epsilon(\Gamma)$ the set of pure ϵ-NE, and by $\mathrm{E}_\epsilon^\diamond(\Gamma)$, the set of mixed ϵ-NE. Note that $\mathrm{E}_\epsilon(\Gamma) \subseteq \mathrm{E}_\epsilon^\diamond(\Gamma)$.

Our main result is that equilibria can be approximated with bounded error, given only a uniform approximation. To present this result, we define the ℓ_∞*-norm* between two compatible games, with the same agents sets $\boldsymbol{P}$ and strategy profile spaces $\boldsymbol{S}$, and with utility functions $\boldsymbol{u}, \boldsymbol{u}'$, respectively, as follows:

$$\left\|\Gamma - \Gamma'\right\|_\infty \doteq \left\|\boldsymbol{u}(\cdot) - \boldsymbol{u}'(\cdot)\right\|_\infty \doteq \sup_{p \in \boldsymbol{P},\, s \in \boldsymbol{S}} |\boldsymbol{u}_p(s) - \boldsymbol{u}'_p(s)| \ .$$

While the ℓ_∞-norm as defined applies only to pure normal-form games, it is in fact sufficient to use this metric even to show that the utilities of mixed strategy profiles approximate one another.

Definition 2.3. Γ' is said to be a **uniform** ϵ**-approximation** of Γ when $\|\Gamma - \Gamma'\|_\infty \leq \epsilon$.

Uniform approximations are so-called because the bound between utility deviations holds *uniformly* over *all* players and strategy profiles. We now present our main result.

Theorem 2.4 (Approximate Equilibria). *If two NFGs, Γ and Γ', are uniform approximations of one another, then:* $\mathrm{E}(\Gamma) \subseteq \mathrm{E}_{2\epsilon}(\Gamma') \subseteq \mathrm{E}_{4\epsilon}(\Gamma)$ *and* $\mathrm{E}^\diamond(\Gamma) \subseteq \mathrm{E}_{2\epsilon}^\diamond(\Gamma') \subseteq \mathrm{E}_{4\epsilon}^\diamond(\Gamma)$.

We have thus established perfect recall by an approximate game, in the sense that the approximate game contains all true positives: i.e., all (exact) equilibria of the original game; and approximately perfect precision, in the sense that all false positives in the approximate game are approximate equilibria in the original game.

3 LEARNING GAMES

Having established perfect recall and approximately perfect precision of uniform approximations, we now tackle the problem of constructing empirical games. For this purpose, we introduce two algorithms. The first, which we call *global sampling* (GS), learns an empirical game from a static sample. This learning is conceptually very simple: simulate the game m times at *all* strategy profiles, and then average the ensuing utilities across simulations. The requisite number of samples, m, is a function of a user-specified desired accuracy ϵ and failure probability δ. Our second algorithm, *progressive sampling with pruning* (PSP), samples dynamically, saving on queries to the simulator at strategy profiles where fewer data are necessary to confidently learn to a desired degree of accuracy.

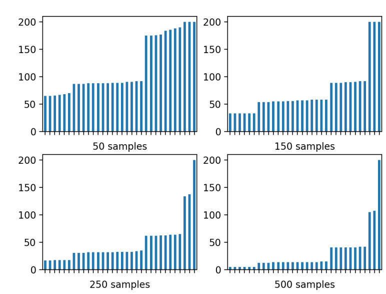
GS and PSP can both be instantiated with various concentration inequalities to obtain uniform convergence guarantees. Building on earlier work [10], we apply Hoeffding's inequality [1]. While doing so requires bounded noise, this is not an inherent limitation of our methodology. We could obtain similar results under varied noise assumptions; e.g., we could assume *subgaussian* or *subexponential* noise, and substitute the appropriate Chernoff bounds.

4 EXPERIMENTAL EVALUATION

In the full version of this paper, we empirically evaluate our algorithms on randomly generated games and finite congestion games [7]. We show that they make frugal use of data, accurately estimate games more often than the theory predicts, and are robust to different forms of noise. We further show that in practice PSP requires significantly fewer data than GS, the current state-of-the-art [10], to produce the same (and often times, better) error rates.

We summarize here one of a suite of experiments we conducted in our evaluation. The goal of this experiment is to evaluate the extent to which Theorem 2.4 holds in practice, when learning empirical games via GS. Figure 1 depicts four histograms, each one plotting the frequencies of strategy profiles deemed pure 2ϵ-Nash equilibria, using exhaustive search to find these equilibria, for four different sample sizes. The profiles not shown had zero frequency. The unique pure Nash equilibrium of the underlying congestion game is correctly identified in all cases for all 200 trials. Moreover, as the number of samples increases, the frequency of false positives (i.e., profiles that are not Nash but are deemed so by GS) decreases.

Figure 1: Frequency of pure 2ϵ-Nash equilibria



5 APPLICATIONS

One important application of our methodology is the estimation of equilibria in so-called *meta-games* [9, 15], which are simplified versions of intractably large games. In a meta-game, instead of modeling every possible strategy an agent might implement, one analyzes a game with a substantially reduced set of strategies, each of which is usually given by a complicated algorithmic procedure (i.e., a heuristic). For example, instead of analyzing a game that models every possible strategy that could be played in a game of Go—a computationally intractable task—one might analyze a reduced version of the game where agents play strategies given by reinforcement learning algorithms [8] (e.g., variants of AlphaGo). Moreover, simulation is in order; and since each run of the game can result in either agent winning (depending on various stochastic elements, including the agents' strategies), one can only obtain noisy utilities. Our techniques are directly applicable to the construction of empirical meta-games, and provide guarantees on the quality of the equilibria of the corresponding simulation-based meta-games.

Another important application of our methodology is in the area of *empirical mechanism design* [13]. In mechanism design, the mechanism (game) designer wishes to optimize the rules of a game so that the ensuing equilibria achieve certain goals. We plan to extend our methodology to learn empirically optimal mechanisms.

ACKNOWLEDGEMENTS

This work was supported by NSF Grant CMMI-1761546.

REFERENCES

[1] Wassily Hoeffding. 1963. Probability inequalities for sums of bounded random variables. *Journal of the American statistical association* 58, 301 (1963), 13–30.

[2] Patrick R Jordan, Christopher Kiekintveld, and Michael P Wellman. 2007. Empirical game-theoretic analysis of the TAC supply chain game. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. ACM, 193.

[3] Patrick R Jordan, Yevgeniy Vorobeychik, and Michael P Wellman. 2008. Searching for approximate equilibria in empirical games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 1063–1070.

[4] Patrick R Jordan and Michael P Wellman. 2010. Designing an ad auctions game for the trading agent competition. In *Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets*. Springer.

[5] Wolfgang Ketter, Markus Peters, and John Collins. 2013. Autonomous Agents in Future Energy Markets: The 2012 Power Trading Agent Competition.. In *AAAI*.

[6] Victor Picheny, Mickael Binois, and Abderrahmane Habbal. 2016. A Bayesian optimization approach to find Nash equilibria. *arXiv preprint arXiv:1611.02440* (2016).

[7] Robert W Rosenthal. 1973. A class of games possessing pure-strategy Nash equilibria. *International Journal of Game Theory* 2, 1 (1973), 65–67.

[8] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. 2016. Mastering the game of Go with deep neural networks and tree search. *nature* 529, 7587 (2016), 484.

[9] Anderson Tavares, Hector Azpurua, Amanda Santos, and Luiz Chaimowicz. 2016. Rock, paper, starcraft: Strategy selection in real-time strategy games. In *The Twelfth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-16)*.

[10] Karl Tuyls, Julien Perolat, Marc Lanctot, Joel Z Leibo, and Thore Graepel. 2018. A Generalised Method for Empirical Game Theoretic Analysis. *arXiv preprint arXiv:1803.06376* (2018).

[11] Leslie G Valiant. 1984. A theory of the learnable. *Commun. ACM* 27, 11 (1984), 1134–1142.

[12] Yevgeniy Vorobeychik. 2010. Probabilistic analysis of simulation-based games. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* 20, 3 (2010), 16.

[13] Yevgeniy Vorobeychik, Christopher Kiekintveld, and Michael P Wellman. 2006. Empirical mechanism design: methods, with application to a supply-chain scenario. In *Proceedings of the 7th ACM conference on Electronic commerce*. ACM.

[14] Yevgeniy Vorobeychik and Michael P Wellman. 2008. Stochastic search methods for Nash equilibrium approximation in simulation-based games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, 1055–1062.

[15] Michael P Wellman. 2006. Methods for empirical game-theoretic analysis. In *AAAI*. 1552–1556.

[16] Michael P Wellman, Tae Hyung Kim, and Quang Duong. 2013. Analyzing incentives for protocol compliance in complex domains: A case study of introduction-based routing. *arXiv preprint arXiv:1306.0388* (2013).