Ishikawa Play

Yoad Lewenberg[*]
yoadlew@cs.huji.ac.il

Zinovi Rabinovich[⋆]
zr@zinovi.net

Jeffrey S. Rosenschein[*]
jeff@cs.huji.ac.il

[*] School of Computer Science and Engineering, The Hebrew University of Jerusalem, Israel
[⋆] Mobileye Vision Technologies Ltd., PO Box 45157, Jerusalem, 97775, Israel

ABSTRACT

Bayes-Nash Equilibrium (BNE) is at the root of many significant applications of modern game theory to multiagent systems, ranging from airport security scheduling, to network analysis, to mechanism design in e-commerce. However, the computational complexity of calculating BNEs makes the process prohibitively costly, and the process does not scale well. On the other hand, finding BNEs by simulating the repeated interaction of adaptive players has been demonstrated to succeed even in very complex domains. Unfortunately, adaptive algorithms that iteratively shift strategy towards an equilibrium (e.g., the Fictitious Play algorithm) do not provide stable performance across all classes of games. Therefore, active research into these stability issues, and the design of new algorithms for interactive BNE calculation, remain highly important.

In this paper we present a variation to the Ishikawa Iteration to calculate a Bayes-Nash Equilibrium. We demonstrate that the Ishikawa algorithm can take an interactive form, which we term Ishikawa Play (I-Play), and apply it in repeated games. Our experimental data shows that variations of the I-Play algorithm are effective in self-play (converge to a BNE), and outperform the Fictitious Play algorithm, while maintaining low computational costs per game cycle.

Categories and Subject Descriptors

I.2 [**Distributed Artificial Intelligence**]: Intelligent agents; I.2 [**Distributed Artificial Intelligence**]: Multiagent systems

General Terms

Algorithms, Experimentation

Keywords

Ishikawa Iteration, Fictitious Play, Equilibria computation

1. INTRODUCTION

Bayes-Nash Equilibrium (BNE) is one of the key solution concepts of modern game theory, and it continues to present both conceptual and computational challenges. Although the significant computational complexity of finding Nash equilibria has been extensively researched (see, for example, [4, 3]), the importance of finding equilibria continues to feed research both into exact and approximate off-line solution methods [17, 15, 10, 11, 14].

However, one conceptual challenge of BNE is that it does not necessarily arise naturally. In particular, if a game is played repeatedly, higher-order dependencies between players and their strategies arise for which standard BNE cannot account. As a result, there have been many new equilibrium concepts introduced in recent years, such as program equilibria [23], cyclic equilibria [24] and sink equilibria [7]. Nonetheless, BNE continues to be an attractive solution for repeated games, for several reasons. First, if the repeated game describes an actual *interactive game play*, the overhead of calculating a BNE is compensated for by the almost zero computational cost incurred during actual game playing, while still providing performance guarantees. Second, several computationally-efficient iterative approximation algorithms exist that can approach a BNE dynamically, during the repeated game interaction itself, thus reducing the overhead of an *a priori* computation of a specific BNE, while maintaining low computational effort per game repetition.

Perhaps the most popular and simplest algorithm of this class is the Fictitious Play (FP) algorithm [2] and its off-line counterpart, the Mann Iteration [16], which was originally developed as a fixed-point computation method. The FP algorithm was initially developed to solve two-player matrix games. Since its inception, however, it has been extended and applied towards practical solutions of such complex games as double [22] and simultaneous [20] auctions, and poker [8]. Although in some classes of games FP provably succeeds [19, 18], it cannot be guaranteed for all classes of games. Furthermore, its rate of convergence towards an equilibrium strategy is known to be low [1]. In addition, the same instability that inhibits FP convergence can also be exploited when used as an actual interactive game player [5].

However, there are fixed-point methods that are faster, and even have better convergence guarantees, than the Mann Iteration (FP's precursor). This suggests that it is possible to improve upon FP performance, with respect to convergence towards equilibrium. In this paper, we explore this topic; specifically, based on the iterative fixed-point method of Ishikawa [12], we compose a series of interactive equilibrium approximation algorithms, which we term *I-Play*, and empirically evaluate their performance.

The rest of the paper is structured as follows. In Section 2 we introduce our repeated game model, and our associated solution and performance measurements. Section 3 formally, yet briefly, reviews iterative equilibrium approximation algorithms based on *better response*. In Section 4 we introduce the Ishikawa Iteration and our I-Play algorithm variations, followed by our experimental evaluation in Section 5. Finally, in Section 6 we conclude, and discuss future work on I-Play algorithms.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013), Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.*

2. MODEL AND MEASUREMENTS

In this paper we adopt the model of a two-player repeated simultaneous-move Bayesian game. A single stage game is defined by the tuple $G = \langle \mathcal{A}^i, u^i, \Omega^i, p^i | i \in \{1,2\} \rangle$ where:

- $\mathcal{A}^i$ is the finite set of actions available to player $i \in \{1,2\}$. We will also denote by $\mathbf{A} = \times_{i \in \{1,2\}} \mathcal{A}^i$ the space of joint actions, and its elements by $(a^i, a^{-i}) = \mathbf{a} \in \mathbf{A}$, where a^i is the action taken by the player i and a^{-i} is the action taken by its opponent.
- Ω^i is the finite set of types that player i can possess.
- $p^i \in \Delta(\Omega^i)$ is the publicly known distribution over Ω^i.
- $u^i(\omega^i, \mathbf{a})$ is the utility player i obtains if it has type $\omega^i \in \Omega^i$ and players take joint action $\mathbf{a} \in \mathbf{A}$.

At every stage $t = 1, \ldots, N$ the game proceeds as follows. First, ω^i is sampled i.i.d. for each player and revealed to the player alone as private information. Then, each player chooses action $a^i \in \mathcal{A}^i$. The joint action $(a^1, a^2) = \mathbf{a} \in \mathbf{A}$ is revealed to both players along with their respective rewards $u^i(\omega^i, \mathbf{a})$. The stage game is played repeatedly an unknown, but finite, number of times (or equivalently, assumed to repeat infinitely).

To complete the game protocol we define the strategy of player $i \in \{1,2\}$ (i.e., how an action is chosen in a single stage game) as a function $\sigma^i : \Omega^i \to \Delta(\mathcal{A}^i)$, where $\sigma^i(\cdot|\omega^i)$ is the distribution over all actions available to the player, given that its private type is ω^i. In other words, we allow players to randomize their choice of actions, so that action $a^i \in \mathcal{A}^i$ is chosen with probability $\sigma^i(a^i|\omega^i)$. We note that the strategy may change over time, and denote the strategy used by player i at stage t by σ^i_t. Similarly, we denote the *joint* strategy at stage t by $\sigma_t = (\sigma^i_t, \sigma^{-i}_t)$. To further facilitate our discussion we define the (expected) utility to a player i of a (joint) strategy σ by

$$\begin{aligned} u^i(\sigma|\omega^i) &= \mathbb{E}\left[u^i(\omega^i, \mathbf{a})\right] \\ &= \sum_{\omega^{-i} \in \Omega^{-i}} \sum_{\mathbf{a} \in \mathbf{A}} u^i(\omega^i, \mathbf{a}) \sigma^i(a^i|\omega^i) \sigma^{-i}(a^{-i}|\omega^{-i}) p^{-i}(\omega^{-i}) \end{aligned}$$

We will also use the expected utility to a player i of a specific action, given the opponent's strategy σ^{-i}:

$$\begin{aligned} &u^i(a^i|\omega^i, \sigma^{-i}) = \\ &\quad \sum_{\omega^{-i}} \sum_{a^{-i} \in \mathcal{A}^{-i}} u^i(\omega^i, (a^i, a^{-i})) \sigma^{-i}(a^{-i}|\omega^{-i}) p^{-i}(\omega^{-i}) \end{aligned}$$

Notice that the single-game stage protocol can be implemented in a decentralized, *interactive* manner, where players are distinct, independently designed and operated entities intent on maximizing their own expected utilities. Given that the game G is played repeatedly, players may strategically change σ^i_t over time in an attempt to increase their (long-term, accumulated) personal gain. While many new concepts have arisen in recent years to describe this high-level behavior [23, 24, 7], the solution concept of *Bayes-Nash Equilibrium* (BNE) continues to be the most popular. This is perhaps due to the fact that players using BNE can reduce their computational load by adhering to a single strategy at all stages of the repeated game.

Definition 1. A joint strategy $\sigma^{BNE} = (\sigma^{i,BNE}, \sigma^{-i,BNE})$ is a Bayes-Nash Equilibrium of a (stage) game G, if for all i, for all $\omega^i \in \Omega^i$ and for all σ^i, it holds that

$$u^i(\sigma^{BNE}|\omega^i) \geq u^i((\sigma^i, \sigma^{-i,BNE})|\omega^i)$$

That is, no player can improve its expected utility for any of its types beyond that defined by σ^{BNE}, if its opponent also does not deviate from $\sigma^{-i,BNE}$.

However, in a repeated game setting, players may not necessarily start *a priori* playing the strategies prescribed by a Bayes-Nash Equilibrium, even if they intend to eventually play a BNE strategy. Rather, players might begin from some ill-informed, random σ^i and σ^{-i}. The players then adapt over time, and eventually either converge to (in a sense, agree on) an equilibrium, or find a more profitable off-equilibrium behavior. An additional complication may come from the fact that different players may adapt differently. This could occur, for example, if instead of using a centralized algorithm (which utilizes repeated play to calculate an equilibrium), *interactive play* is used in a decentralized system where players are independent and employ different algorithms.

To measure whether players are better off or worse off while adapting and playing off-equilibrium strategies, we introduce *normalized utility* for games with a unique BNE.

Definition 2. Let σ^{BNE} be the unique Bayes-Nash equilibrium of a stage game G. We define μ^i, the *total* expected utility of player i as $\mu^i = \mathbb{E}_{\omega^i}[u^i(\sigma^{BNE}|\omega^i)]$. Similarly the *total* variance of player i's utility is

$$\Sigma^i = Var_{\omega^i \sim p^i, \omega^{-i} \sim p^{-i}, a \sim \sigma^{BNE}(\omega^i, \omega^{-i})} \left[u^i(\omega^i, a)\right]$$

Finally, we define the *empirical normalized average utility* of player i,

$$\Phi^i_{[1:N]} = \frac{u^i_{[1:N]} - \mu^i}{\sqrt{\Sigma^i}},$$

where $u^i_{[1:N]}$ is the empirical average utility given by $u^i_{[1:N]} = \frac{1}{N} \sum_{t=1}^{N} u^i(\omega^i_t, \mathbf{a}_t)$ with $\mathbf{a}_t$ and ω^i_t being the joint action and player i's type during the stage t game.

Notice that the normalized average utility is, in a sense, a scale-free measure of performance due to the following lemma.

LEMMA 1. *Let there be two (single stage) games $G = \langle \mathcal{A}^i, u^i, \Omega^i, p^i | i \in \{1,2\} \rangle$ and $G' = \langle \mathcal{A}^i, \widehat{u}^i, \Omega^i, p^i | i \in \{1,2\} \rangle$, so that $\widehat{u}^i = cu^i + b$ with $c, b \in \Re$ and $c > 0$. Assume that both G and G' have been played repeatedly N times. Then for any sequence of joint actions, $\mathbf{a}_{[1:N]}$, and for any sequence of private types of player i, $\omega^i_{[1:N]}$, it holds that $\Phi^i_{[1:N]} = \widehat{\Phi}^i_{[1:N]}$, where $\Phi^i_{[1:N]}$ and $\widehat{\Phi}^i_{[1:N]}$ are the normalized average utilities of game G and G' respectively.*

In other words, the normalized average utility, Φ, is invariant under linear transformations of the stage game utility function.

Similarly, it can be shown[1] that $\Phi^i_{[1:N]}$ has the following properties:

- if both players apply the equilibrium strategy, then $\Phi^i_{[1:N]}$ tends to zero as the number of played stages, N, grows;
- if player i's opponent was playing an off-equilibrium strategy and player i was persistently able to exploit that fact, then $\Phi^i_{[1:N]} > 0$.

[1] We omit the proof of the lemma and associated conclusions, due to their length, tedious technicality, and ease of reproduction.

Although non-zero Φ^i implies that a non-equilibrium strategy is being played, it may still perform sufficiently close to an equilibrium. In fact, utility-based proximity to an equilibrium strategy is termed an ϵ-equilibrium and is defined as follows.

Definition 3. A joint strategy $\sigma^{\epsilon-BNE}$ is an ϵ Bayes-Nash Equilibrium (ϵ-BNE) of a (stage) game G, if for all i, for all $\omega^i \in \Omega^i$ and for all σ^i it holds that

$$u^i(\sigma^{\epsilon-BNE}|\omega^i) + \epsilon \geq u^i((\sigma^i, \sigma^{-i,\epsilon-BNE})|\omega^i)$$

In other words, the best possible deviation from $\sigma^{\epsilon-BNE}$ cannot give an increase of more than ϵ utility. The margin of error ϵ becomes a good measure of how far away the joint strategy σ is from being an equilibrium. We therefore define the distance $\Psi_\epsilon(\sigma)$ as a measure of "how much better off" any player i can be by deviating from σ^i at some type $\omega^i \in \Omega^i$:

Definition 4. Let σ be a joint-policy candidate to be an ϵ-BNE. The utility discrepancy bound of this joint policy, Ψ_ϵ is given by

$$\Psi_\epsilon(\sigma) = \max_{i\in\{1,2\}} \max_{\omega^i\in\Omega^i} \max_{a^i\in\mathcal{A}^i} u^i(a^i|\omega^i, \sigma^{-i}) - u^i(\sigma^i|\omega^i, \sigma^{-i})$$

3. ITERATIVE APPROXIMATIONS

An equilibrium is by definition a stable point of maximization of the expected utility function, and as such can sometimes be solved by iterative approximation methods. These methods start from an arbitrary joint strategy, and then iteratively improve it in some sense, until the next iteration can no longer produce a significant change in the strategy or its expected utility, thus producing an approximate BNE equilibrium strategy. Two widely-applied methods of this class are Best (or Better) Response (BR), and the Mann Iteration (MI) algorithms. While Best Response will serve as a tool in defining the main algorithmic contribution of this paper, Ishikawa Play (I-Play), the interactive version of the MI algorithm, will serve as a baseline for I-Play's performance tests.

3.1 Best and Better Responses

Best/Better Response (BR) methods repeatedly apply a transformation $\sigma \mapsto T(\sigma) = (\sigma^{i,BR}, \sigma^{-i,BR})$ so that the unilateral utility gain is nonnegative. That is, for all ω^i (respectively ω^{-i}) it holds that $u^i((\sigma^{i,BR}, \sigma^{-i})|\omega^i) - u^i(\sigma, \omega^i) \geq 0$ (respectively, $u^{-i}((\sigma^i, \sigma^{-i,BR})|\omega^{-i}) - u^{-i}(\sigma, \omega^{-i}) \geq 0$). The better response transformation T is called the *best response* transformation if utility gain is maximized for all types.

Notice that if the iteration $\sigma_t = T(\sigma_{t-1})$ converges, then σ_t is an ϵ-equilibrium, and if the convergence is exact, i.e., $\sigma_t = T(\sigma_{t-1}) = \sigma_{t-1}$, then $\sigma_t = \sigma^{BNE}$ is the equilibrium strategy by definition.

In this paper we will distinguish between three improvement transformations T:

- T_1, *best response* transformation with tie-breaking;
- T_2, *smoothed best response* parameterized by $\lambda > 0$;
- T_3, *continuous better response*.

3.1.1 Best Response, T_1

Let us enumerate all actions in $\mathcal{A}^i = \{a_1, ..., a_k\}$, and associate with each action a_j a unit vector $e_j \in \Delta(\mathcal{A}^i) \subset \Re^k$; then the set of best responses of player i having type $\omega^i \in \Omega^i$ to an opponent strategy σ^{-i} is given by

$$BR^i(\omega^i|\sigma^{-i}) = \{e_j | u^i(\mathcal{A}^i|\omega^i, \sigma^{-i}) \leq u^i(a_j|\omega^i, \sigma^{-i})\}.$$

That is, it is the set of all actions by which player i maximizes its utility for a given type and strategy of its opponent. Then for a given σ we define $(\widehat{\sigma}^i, \widehat{\sigma}^{-i}) = \widehat{\sigma} = T_1(\sigma)$ as follows:

$$\widehat{\sigma}^i(\cdot|\omega^i) = \frac{1}{|BR^i(\omega^i|\sigma^{-i})|} \sum_{e_j\in BR^i(\omega^i|\sigma^{-i})} e_j$$

In other words, player i selects with equal probability all those actions that provide the best possible expected utility against the given strategy of its opponent, and the opponent player $-i$ acts similarly. Notice also that $\sigma^{i,BNE}(\cdot|\omega^i)$ is necessarily a convex combination of vectors in $BR^i(\omega^i|\sigma^{-i,BNE})$, and that all such combinations yield the same expected utility.

3.1.2 Smoothed Best Response, T_2

Instead of concentrating only on the best possible actions against a given opponent strategy, as did T_1, T_2 *prefers* actions that yield higher utility. Formally, for $(\widehat{\sigma}^i, \widehat{\sigma}^{-i}) = \widehat{\sigma} = T_2(\sigma)$ it holds that

$$\widehat{\sigma}^i(a^i|\omega^i) \propto \exp(\frac{1}{\lambda} u^i(a^i|\omega^i, \sigma^{-i}))$$

Notice that as λ tends to zero, the outcome of T_2 approaches that of T_1, i.e., *smoothed best response* tends to the *best response* of a player. This response function is also known as the Logit Quantal Response, among other names.

3.1.3 Continuous Better Response, T_3

Continuous better response concentrates on the relative utility gain of choosing an action, rather than on the expected reward that the action produces. More formally, let

$$\delta^i(a^i|\omega^i, \sigma) = \max\{0, u^i(a^i|\omega^i, \sigma^{-i}) - u^i(\sigma|\omega^i)\}.$$

Then for $(\widehat{\sigma}^i, \widehat{\sigma}^{-i}) = \widehat{\sigma} = T_3(\sigma)$ it holds that

$$\widehat{\sigma}^i(a^i|\omega^i) = \frac{\sigma^i(a^i|\omega^i) + \delta^i(a^i|\omega^i, \sigma)}{1 + \sum_a \delta^i(a|\omega^i, \sigma)}$$

Interestingly, the Continuous better response function was originally used by Nash to prove equilibrium existence via Brouwer's fixed-point theorem.

The following properties of T_1, T_2, T_3 should be noted. $T_1(\sigma)$ and $T_2(\sigma)$ "snap" to the best actions in response to opponent strategy, while $T_3(\sigma)$ partially preserves the original behavior of the argument strategy, σ. Another property worthy of note is that T_1 and T_2 do not necessarily preserve the equilibrium strategy, i.e., there's no guarantee that $T_1(\sigma^{BNE}) = \sigma^{BNE}$, nor that $T_2(\sigma^{BNE}) = \sigma^{BNE}$. On the other hand, T_3 does have the equilibrium preservation property, and it holds that $T_3(\sigma^{BNE}) = \sigma^{BNE}$.

Now, it is known that BR iteration converges in some specific classes of games, and some efficiency results are known (see, e.g., [6] and the references therein). However, some negative results also exist, especially for more complex games [9], which encourage the use of other, more intricate iterative methods for BNE calculation. One such method is the Mann Iteration and its interactive version, Fictitious Play, which we describe next.

3.2 Mann Iteration and Fictitious Play

Mann Iteration (MI) was introduced [16] to calculate a fixed point of a function—in our case, it would be the fixed point of the Best Response transformation, T_1. Unlike the pure iterative application of T_1, MI does not simply augment the given strategy, but rather aggregates such augmentations over time. To do so MI relies on an internal memory that holds a *belief* regarding the most recent approximation of the equilibrium strategy. At each iteration, MI calculates the best response of each player to the *belief*, and then merges the two together.

Set x_0 arbitrarily to a joint strategy
Set the equilibrium precision ϵ, and $t = 0$
repeat
  $x_{t+1} = (1 - \alpha_t)\, x_t + \alpha_t T\,(x_t)$
  $t \leftarrow t + 1$
until x_t is ϵ-equilibrium

Figure 1: Mann Iteration

Formally, the beliefs held by MI have the same form as a joint strategy, and we will use the same notational tools for MI beliefs and strategies. Denote by $x_t = (x_t^i, x_t^{-i})$ MI beliefs at iteration t; then $x_{t+1} = (1 - \alpha_t)x_t + \alpha_t T_1(x_t)$, where T_1 is the best response transformation defined in Section 3.1 (see Figure 1), and α_t is the sequence of coefficients so that (i) $\lim_{n \to \infty} \alpha_n = 0$ and (ii) $\sum_{t=1}^{\infty} \alpha_t = \infty$. To satisfy coefficient requirements, α_t is commonly set to $\frac{1}{t+1}$. Similarly to the BR iteration, MI stops whenever x_{t+1} is an equilibrium strategy or the utility gains of the best response to x_{t+1} are negligible. In fact, it is easy to show that if x_t converges, it converges to σ^{BNE} of the game to which MI is applied.

Although Mann Iteration has been studied for many years as a means of fixed point calculation, applied to equilibrium calculations it is better known by a different name: Fictitious Play (FP). Fictitious Play is the interactive version of Mann Iteration, and was introduced as a means of calculating a Nash equilibrium in two-player matrix games by Brown [2]. Since that initial publication, FP has been extended to apply to many game classes, including extremely complex double auctions [22].

1: Set arbitrary joint strategy x_0
2: Set $t = 0$
3: **repeat**
4:   Receive ω^i
5:   Sample $a_t^i \sim x_t^i(\cdot|\omega^i)$
6:   Report a_t^i
7:   Receive implemented (a_t^i, a_t^{-i}) report and $u^i(\omega^i, (a^i, a^{-i}))$
8:   Let $e_t^{-i} \in \Delta(\mathcal{A}^{-i})$, where $e_t^{-i}(a) = 1 \iff a = a_t^{-i}$
9:   Let $\alpha_t = \frac{1}{t+1}$
10:   Set $\forall \omega^{-i} \in \Omega^{-i}$ $x_{t+1}^{-i}(\cdot|\omega^{-i}) = (1 - \alpha_t)x_t^{-i}(\cdot|\omega^{-i}) + \alpha_t e_t$,
11:   Let $x = T_1(x_t^i, x_{t+1}^{-i})$, and set $x_{t+1}^i = x^i$
12:   Set $t = t + 1$
13: **until** game terminated

Figure 2: Fictitious Play (interactive Mann Iteration)

FP is obtained from Mann Iteration by distribution of the calculation between the players, as is shown in Figure 2. Notice that it exactly corresponds to Brown's definition of FP: continually play the best response to the empirical frequency of opponent actions.[2] To understand why it is indeed an MI implementation, notice that MI can be broken into two steps: (i) calculation of T_1, and (ii) weighted aggregation of the outcome. FP calculates the best response, T_1, of the player (line 11), but aggregates the actions of the opponent (line 10). When run simultaneously by all players, a complete MI calculation is performed.

[2]Brown originally intuited this by adding that each player assumes that the other plays a fixed strategy, and therefore best response to the empirical frequency is optimal.

4. ISHIKAWA ITERATION AND PLAY

Although Fictitious Play is a provably effective means of obtaining an equilibrium in several interesting settings [19, 18], its convergence is unstable, and its convergence rate is low [1]. We therefore propose using another fixed point iterative method, specifically the Ishikawa Iteration [12], as a means of calculating a BNE.

1: Set x_0 arbitrarily to a joint strategy
2: Set the equilibrium precision ϵ, and $t = 0$
3: **repeat**
4:   $y_t = (1 - \beta_t)\, x_t + \beta_t T\,(x_t)$
5:   $x_{t+1} = (1 - \alpha_t)\, x_t + \alpha_t T\,(y_t)$
6:   $t \leftarrow t + 1$
7: **until** x_t is ϵ-equilibrium

Figure 3: Ishikawa Iteration

Ishikawa Iteration (shown in Figure 3) is a two-step process that can be described as a Mann Iteration with one-step lookahead. Recall that Mann Iteration directly mixes its beliefs, x_t, with a better response to them, $T(x_t)$. Instead, Ishikawa Iteration considers the effect such a mixture would have on beliefs (line 4), and repeats the policy improvement step (line 5). This also necessitates a gentler treatment of the mixture coefficients. Ishikawa requirements for α_t and β_t are: (i) $0 \leq \alpha_t \leq \beta_t \leq 1$; (ii) $\lim_{t \to \infty} \beta_t = 0$; (iii) $\sum_{t=1}^{\infty} \alpha_t \beta_t = \infty$. The simplest choice of coefficients to satisfy these conditions is $\alpha_t = \beta_t = \frac{1}{\sqrt{t+1}}$, and we adopt these coefficient sequences in our experiments.

Now, given a stage game G, Ishikawa Iteration can always be applied to attempt a BNE calculation, just as with the Mann Iteration. However, its operation contains a key assumption that prevents it from being directly applied in the game play, namely that the strategy of both players is under centralized control of the algorithm. In the following sections, we augment and distribute the computation of the Ishikawa Iteration to obtain a series of interactive algorithms that we term Ishikawa Play (I-Play). The distribution method echoes that used to obtain Fictitious Play from the Mann Iteration. We experimentally demonstrate the efficacy of I-Play variations in interactive repeated game playing in Section 5.

4.1 I-Play Variations

Figure 4 gives the distributed version of the Ishikawa Iteration to allow for interactive game play: I-Play-0. I-Play-0 implicitly includes several variations through the introduction of the *play rule* $p_t = T^{PR}(x_t, y_{t-1}, \omega^i)$. Though only one of them gives a complete distributed implementation of the Ishikawa Iteration (specifically, T^{PR}_{Tx}), we have experimented with several such play rules T^{PR}, where T denotes any of the better response functions from Section 3.1:

- $T^{PR}_x(x_t, y_{t-1}, \omega^i) = x_t^i(\cdot|\omega^i)$
- $T^{PR}_y(x_t, y_{t-1}, \omega^i) = y_t^i(\cdot|\omega^i)$
- $T^{PR}_{Tx}(x_t, y_{t-1}, \omega^i) = (T(x_t))^i(\cdot|\omega^i)$
- $T^{PR}_{Ty}(x_t, y_{t-1}, \omega^i) = (T(y_{t-1}))^i(\cdot|\omega^i)$

We have also tested a variation that uses a weighted combination of FP and I-Play-0 beliefs, I-Play-γ, as is shown in Figure 5. For

Set x_0 arbitrary joint strategy
Set $t = 0$
repeat
  Receive ω^i
  if t=0 **then**
    Set a^i_t arbitrarily
  else
    Calculate play rule $p_t = T^{PR}(x_t, y_{t-1}, \omega^i) \in \Delta(\mathcal{A}^i)$
    Sample $a^i_t \sim p_t$
  end if
  Report a^i_t
  Receive (a^i_t, a^{-i}_t) report and $u^i(\omega^i, (a^i_t, a^{-i}_t))$
  Let $e^{-i}_t \in \Delta(\mathcal{A}^{-i})$, where $e^{-i}_t(a) = 1 \iff a = a^{-i}_t$
  Let $e^i_t \in \Delta(\mathcal{A}^i)$, where $e^i_t(a) = 1 \iff a = a^i_t$
  Let $\alpha_t = \beta_t = \frac{1}{\sqrt{t+1}}$
  Let $y^{-i}_t(\cdot|\omega^{-i}) = (1-\beta_t)\, x^{-i}_t(\cdot|\omega^{-i}) + \beta_t e^{-i}_t$
  Let $y^i_t(\cdot|\omega^i) = (1-\beta_t)\, x^i_t(\cdot|\omega^i) + \beta_t e^i_t$
  Set $x_{t+1} = (1-\alpha_t)\, x_t + \alpha_t T\,(y_t)$
  Set $t = t + 1$
until game is terminated

Figure 4: I-Play-0

brevity, we denote by ψ the empirical frequencies of actions chosen by the players, and omit stage indexing. We also note that γ is an externally-set algorithm parameter.

Set arbitrary joint strategy x_0
Set $\psi = (\psi^i, \psi^{-i}) = (\vec{0}, \vec{0})$
Set $t = 0$
repeat
  Receive ω^i
  if t=0 **then**
    Set a^i_t arbitrarily
  else
    Calculate play rule $p_t = T^{PR}(x_t, y_{t-1}, \omega^i) \in \Delta(\mathcal{A}^i)$
    Sample $a^i_t \sim p_t$
  end if
  Report a^i_t
  Receive (a^i_t, a^{-i}_t) report and $u^i(\omega^i, (a^i_t, a^{-i}_t))$
  Let $e^{-i}_t \in \Delta(\mathcal{A}^{-i})$, where $e^{-i}_t(a) = 1 \iff a = a^{-i}_t$
  Let $e^i_t \in \Delta(\mathcal{A}^i)$, where $e^i_t(a) = 1 \iff a = a^i_t$
  Let $\alpha_t = \beta_t = \frac{1}{\sqrt{t+1}}$
  Set $\psi = (1 - \frac{1}{t+1})\psi + \frac{1}{t+1}(e^i_t, e^{-i}_t)$
  Let $\widehat{x}^i_t(\cdot|\omega^i) = (1-\gamma)\, x^i_t(\cdot|\omega^i) + \gamma\psi^i$
  Let $\widehat{x}^{-i}_t(\cdot|\omega^{-i}) = (1-\gamma)\, x^{-i}_t(\cdot|\omega^{-i}) + \gamma\psi^{-i}$
  Let $y_t = (1-\beta_t)\widehat{x}_t + \beta_t T(\widehat{x}_t)$
  Let $x_{t+1} = (1-\alpha_t) y_t + \alpha_t T(y_t)$
  Set $t = t + 1$
until game is terminated

Figure 5: I-Play-γ

5. EXPERIMENTAL SETUP AND RESULTS

In our experiments we concentrated on games with a unique Bayes-Nash Equilibrium. We started by creating a library of random games. We first generated random payoff functions by uniformly sampling values from the discrete range of [0 : 99], then verified each resulting game as having a unique BNE using the Gamut/Gambit suite. The library contained 10,000 normal form games (degenerate Bayesian games with a single type) with action set sizes ranging from 2 to 5 actions. The library also contained 5000 Bayesian games with 2 types and 2 actions per player. For each of our game structures we thus obtained 300 random games, and ran each 100 times for 5000 iterations for each combination of player algorithms to create statistically sufficient data.

Our experimental data shows that Ishikawa Iteration and I-Play variations were effective both as a means of discovering a BNE and as a reasonable means of interactive play.[3] Below we review the collected empirical data in more detail, grouping various experiments by the game structure, interactive vs. off-line play, and play rules used by the algorithms.

5.1 Normal form games

We viewed normal form games as degenerate Bayesian games with $|\Omega^i| = 1$. However, we allowed for larger action spaces to allow for non-trivial cyclic behavior observed in adaptive algorithms (see, e.g., [21]). First we note that Ishikawa Iteration quickly converges to an ϵ-BNE with ever-decreasing Ψ_ϵ. Each panel of Figure 6 shows $\Psi_\epsilon(x_t)$ as a function of time for each one of the possible belief improvement transformations, T_k, and clearly demonstrates continual improvement of the Ishikawa beliefs as an ϵ-BNE. Although for larger games performance naturally deteriorates, since the space of possible policies dramatically increases in size, Ishikawa Iteration continues to perform well. In fact, as Figure 7 shows, Ishikawa Iteration outperforms the MI/FP method, and more so for larger games.

Having confirmed that Ishikawa Iteration appears to be effective, and outperforms its natural baseline of MI/FP as a method of computing ϵ-BNE, we proceeded to test the interactive properties of Ishikawa variations I-Play-0 and I-Play-γ.

Interestingly, the naive distribution of the Ishikawa Iteration into an interactive form that produced I-Play-0, while tending towards the empirical utility of a BNE, did not exploit the other player. In contrast, FP and I-Play-γ successfully managed to exploit off-equilibrium play. Figure 8 shows the development of the empirical normalized average utility Φ over time for two baseline algorithms (fixed distribution, and FP) when playing against FP or I-Play variations. The data also shows that, in most cases, I-Play-γ is more flexible and capable of exploiting its opponents than FP.

5.2 Bayesian games

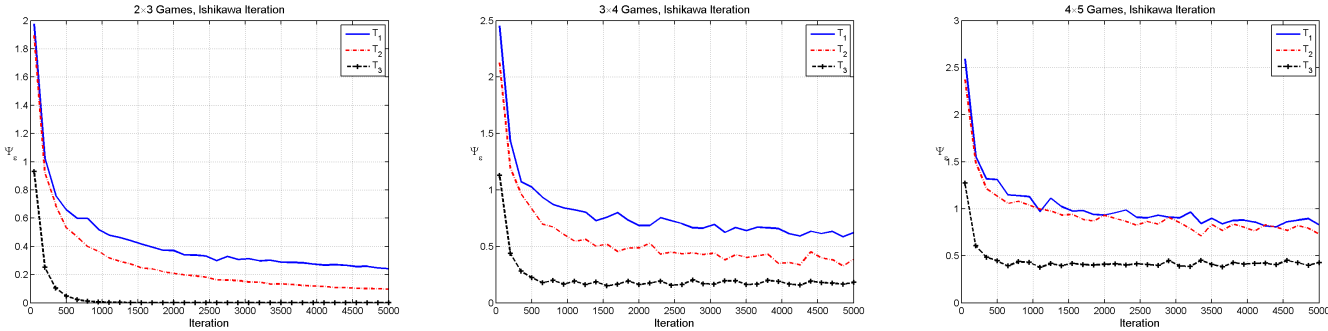
Since we formulated our I-Play variations for BNE, we also performed a set of tests with non-degenerate Bayesian games. We were particularly interested in reconfirming two properties indicated by our data in the normal form games; specifically, that

- Continuous Better Response, T_3, would in practice create the best algorithmic variation of Ishikawa Iteration (originally indicated by the data in Figure 6);
- Ishikawa and I-Play would significantly outperform MI/FP in Bayesian games, since in larger and more complex games the difference is more pronounced (suggested by the data in Figure 7).

Indeed, as Figure 9 shows, Continuous Better Response (T_3) and Smoothed Best Response (T_2) easily outperformed the standard Best Response (T_1) update. In fact, the convergence time almost halved with every step from T_1 to T_3.

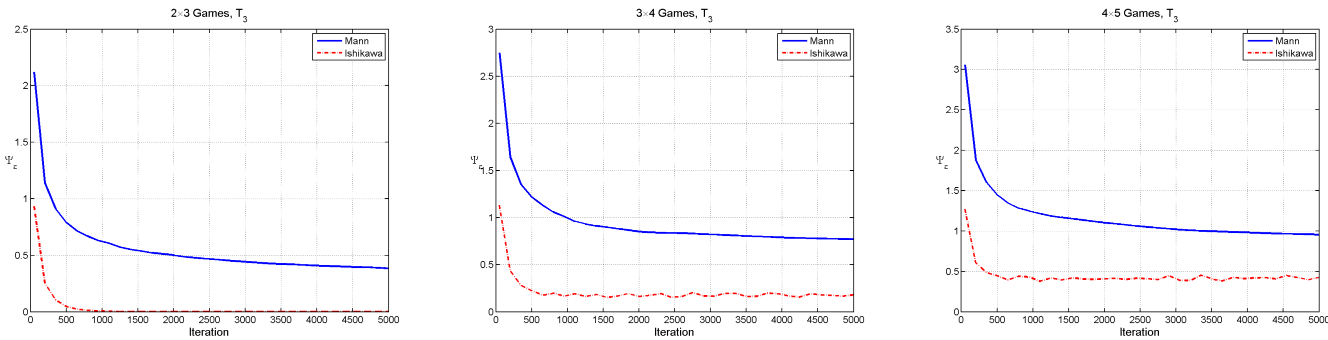
Our second conjecture had mixed results. On the one hand, MI/FP outperformed Ishikawa Iteration when the Best Response,

[3]Notice again that we assume algorithms are tasked to agree on and play an equilibrium-like strategy in the long run.



(a) 2×3 Games (b) 3×4 Games (c) 4×5 Games

Figure 6: Ishikawa Iteration: Ψ_ϵ convergence in normal form games



(a) 2×3 Games (b) 3×4 Games (c) 4×5 Games

Figure 7: Ishikawa Iteration vs. Mann Iteration: Ψ_ϵ convergence in normal form games, T_3 belief update.

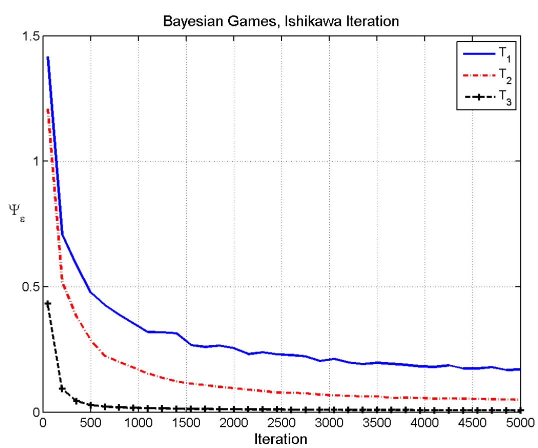


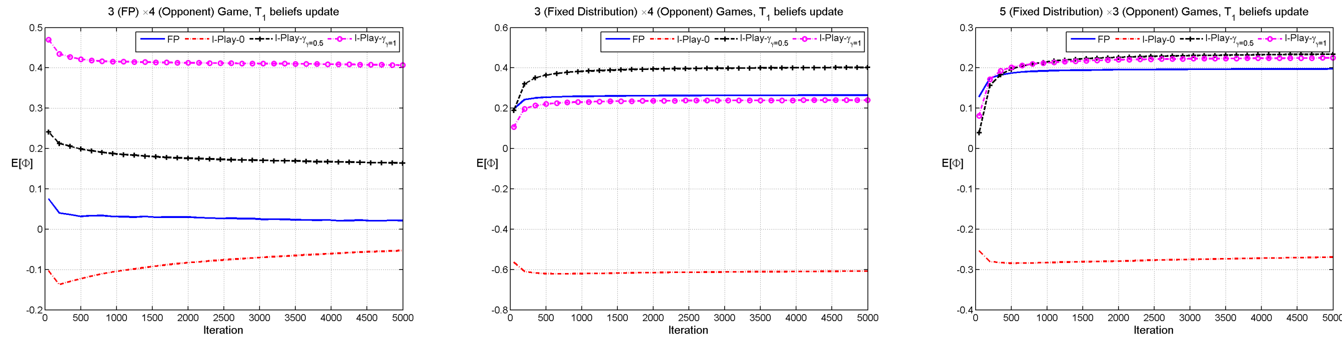
Figure 9: Ishikawa Iteration in Bayesian Games: ϵ-BNE quality, Ψ_ϵ

T_1, update was used. However, with the introduction of the Smoothed Best Response, T_2, and Continuous Better Response, T_3, Ishikawa Iteration turned the tables on MI/FP. Figure 10 shows this outcome, and in particular demonstrates the staggering difference in favor of Ishikawa Iteration with T_3.

We also measured the explicit distance between the strategy produced by an algorithm and the actual equilibrium strategy. Figure 11 shows that not only the utility-based proximity, ϵ-BNE, was achieved, but that the strategies themselves converged as well. In fact, the ϵ-BNE quality and the discrepancy in the strategy space mirrored each other's behavior. In particular, Ishikawa outperformed MI/FP by far when combined with the T_3 update. I-Play variations also performed well in competitive situations. For example, Figure 12 shows the development of the (average) normalized utility of various algorithms when matched with the interactive Mann Iteration (Fictitious Play) opponent using best response functions T_1, T_2 and T_3, respectively. As with our data for normal form games, here too I-Play-0 (distributed Ishikawa in its purest form, i.e., with best response, T_1, update), behaved poorly against FP and was exploited by the latter. In contrast, I-Play-γ successfully outwitted FP. Even though, when T_2 and T_3 were used (Figures 12(b) and 12(c) respectively), the eventual payoff showed convergence to that of the equilibrium strategy, the off-equilibrium play was effectively exploited. Surprisingly, I-Play-0 combined with Continuous Better Response, T_3 improves its relative performance.

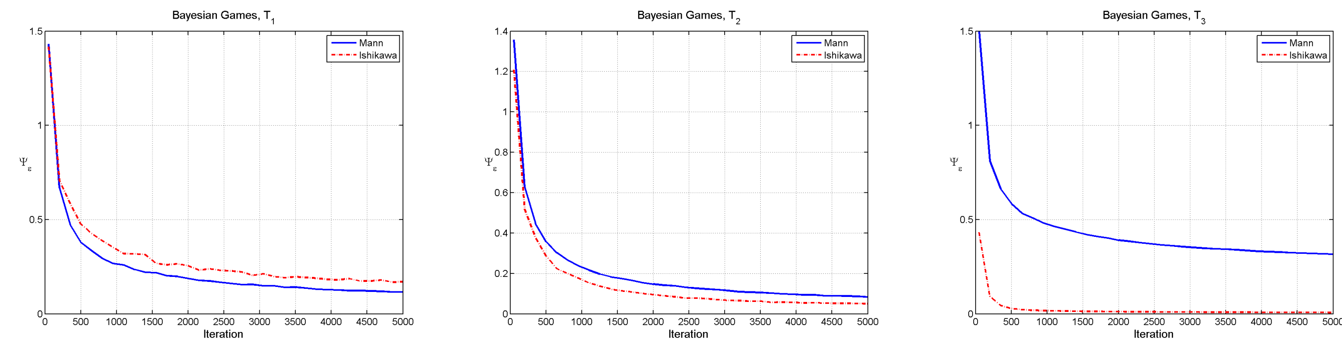
6. CONCLUSIONS AND FUTURE WORK

Iterative methods of Bayes-Nash (BNE) calculation are at today's forefront of the more complex BNE applications, both as a component of a larger algorithmic structure [8] and as a stand-alone algorithm [20]. Unfortunately, adaptive algorithms that iteratively shift strategy towards an equilibrium (e.g., the Fictitious Play algorithm) do not provide stable performance across all classes of games. As a result, a real need arises to develop new iterative BNE calculation algorithms.



(a) 3×4 Games, MI/FP (b) 3×4 Games, Fixed Distribution (c) 5×3 Games, Fixed Distribution

Figure 8: Normal form games: Interactive performance of Fictitious Play (FP) and I-Play variations (I-Play-0 with play rule T_{Tx}^{PR}; I-Play-γ with play rule T_{Ty}^{PR}) matched with Fictious Play and Fixed Distribution; T_1 belief update



(a) Best Response, T_1 (b) Smoothed Best Response, T_2 (c) Continuous Better Response, T_3

Figure 10: Bayesian Games: Ishikawa vs. Mann Iteration: ϵ-BNE quality, Ψ_ϵ

In this paper we took another step in this direction, introducing a set of novel iterative equilibrium approximation algorithms: I-Play-0 and I-Play-γ. The algorithms are based on distribution of the Ishikawa Iteration method for fixed point calculation. Since interactive properties of such distributed calculations are notoriously difficult to analyze formally, we experimentally supported the validity of our algorithms, showing their relative efficiency to another iterative approximation algorithm, Fictitious Play. Nonetheless, the formal analysis of our algorithms is a part of our future work.

Although Ishikawa Iteration has been extensively studied for its convergence for various non-expansive mappings (going as far back as, e.g., [13]), its study with respect to game theoretic equilibria has been lacking. Future work could remedy this shortcoming by formally analyzing Ishikawa Iteration and I-Play, seeking to identify classes of games where I-Play is guaranteed to converge to a Bayes-Nash Equilibrium. In particular, we would like to investigate equilibrium selection properties of our algorithm in games where multiple BNEs are present. In the light of research by de Cote et al. [5], it would also be interesting to test the abilities of I-Play to resist manipulation by higher-order algorithms.

Finally, it would be worthwhile to apply I-Play at a larger scale and in more applicative domains, analogously to what has been done for the Fictitious Play algorithm [22].

Acknowledgments

This work was supported by Israel Science Foundation grant #1227/12, Israel Ministry of Science and Technology grant #3-6797, and the Google Inter-University Center for Electronic Markets and Auctions.

7. REFERENCES

[1] F. Brandt, F. Fischer, and P. Harrenstein. On the rate of convergence of fictitious play. In *Algorithmic Game Theory*, volume 6386 of *LNCS*, pages 102–113. Springer, 2010.

[2] G. W. Brown. Iterative solutions of games by fictitious play. In *Activity Analysis of Production and Allocation*. Wiley, 1951.

[3] V. Conitzer and T. Sandholm. New complexity results about Nash equilibria. *Games and Economic Behavior*, 63(2):621–641, 2008.

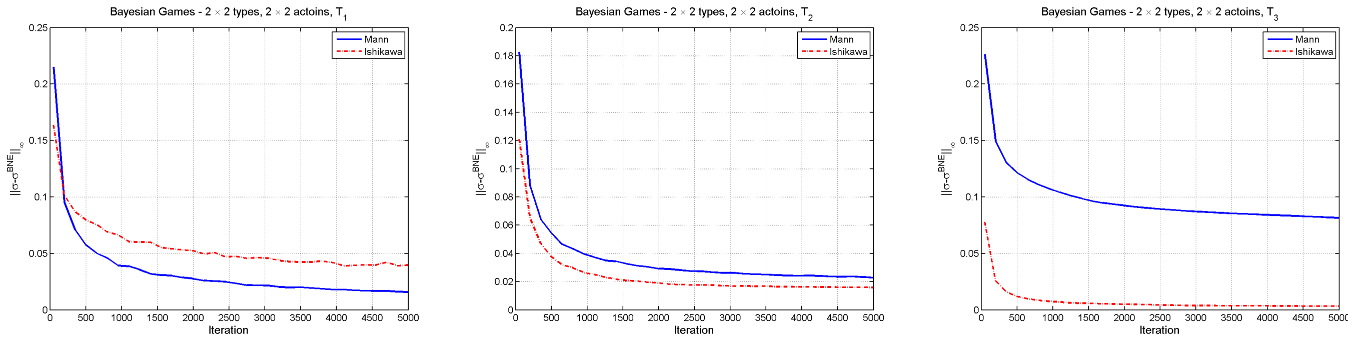
[4] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. In *Proceedings of the 38th ACM STOC*, pages 71–78, 2006.

[5] E. M. de Cote and N. R. Jennings. Planning against fictitious players in repeated normal form games. In *Proceedings of the 9th AAMAS*, pages 1073–1080, 2010.

[6] A. Epstein, B. Awerbuch, Y. Azar, V. Mirrokni, and A. Skopalik. Fast convergence to nearly optimal solutions in potential games. In *Proceedings of the 9th ACM EC*, 2008.

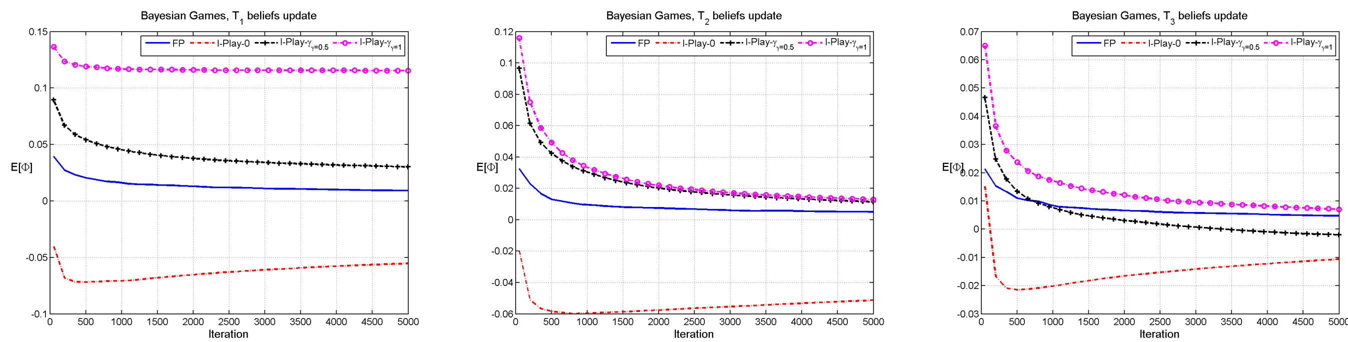
[7] A. Fabrikant and C. Papadimitriou. The complexity of game dynamics: BGP oscillations, sink equilibria, and beyond. In *Proceedings of SODA*, pages 844–853, 2008.

[8] S. Ganzfried and T. Sandholm. Computing an approximation jam/fold equilibrium for 3-player no-limit Texas hold'em tournaments. In *Proceedings of 7th AAMAS*, 2008.



(a) Best Response, T_1 (b) Smoothed Best Response, T_2 (c) Continuous Better Response, T_3

Figure 11: Bayesian Games: Ishikawa vs. Mann Iteration: strategy space proximity to σ^{BNE}



(a) Best Response, T_1 (b) Smoothed Best Response, T_2 (c) Continuous Better Response, T_3

Figure 12: Bayesian Games: Fictitious Play (FP) and I-Play variations (I-Play-0 with play rule T^{PR}_{Tx}; I-Play-γ with play rule T^{PR}_{Ty}) matched with Fictitious Play: Normalized Utility

[9] E. H. Gerding, R. K. Dash, D. C. K. Yuen, and N. R. Jennings. Bidding optimally in concurrent second-price auctions of perfectly substitutable goods. In *Sixth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 267–274, 2007.

[10] A. Gilpin, S. Hoda, J. Peòa, and T. Sandholm. Gradient-based algorithms for finding Nash equilibria in extensive form games. In *Proceedings of the 3rd WINE*, pages 57–69, Berlin, Heidelberg, 2007. Springer-Verlag.

[11] S. Hoda, A. Gilpin, J. Peòa, and T. Sandholm. Smoothing techniques for computing Nash equilibria of sequential games. *Math. Oper. Res.*, 35(2):494–512, May 2010.

[12] S. Ishikawa. Fixed points by a new iteration method. *Proceedings of the American Mathematical Society (AMS)*, 44(1):147–150, 1974.

[13] S. Ishikawa. Fixed points and iteration of a non-expansive mapping in a Banach space. *Proceedings of the American Mathematical Society (AMS)*, 59(1):65–71, 1976.

[14] P. R. Jordan, Y. Vorobeychik, and M. P. Wellman. Searching for approximate equilibria in empirical games. In *Proceedings of AAMAS*, pages 1063–1070, 2008.

[15] D. Koller, N. Megiddo, and B. von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14:247–259, 1996.

[16] W. R. Mann. Mean value methods in iteration. *Proceedings of the American Mathematical Society*, 4:506–510, 1953.

[17] R. D. McKelvey, A. M. McLennan, and T. L. Turocy. *GAMBIT: Software Tools for Game Theory*, 2010.

[18] D. Monderer and L. S. Shapley. Fictitious play property for games with identical interests. *Journal of Economic Theory*, 68(1):258–265, 1996.

[19] D. Monderer and L. S. Shapley. Potential games. *Games and Economic Behavior*, 14:124–143, 1996.

[20] Z. Rabinovich, E. H. Gerding, M. Polukarov, and N. Jennings. Generalised fictitious play for a continuum of anonymous players. In *Proceedings of the 21st IJCAI*, pages 245–250, July 2009.

[21] L. S. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, pages 1–28. Princeton University Press, 1964.

[22] B. Shi, E. Gerding, P. Vytelingum, and N. R. Jennings. An equilibrium analysis of competing double auction marketplaces using fictitious play. In *Proceedings of the 19th ECAI*, pages 575–580, 2010.

[23] M. Tennenholtz. Program equilibrium. *Games and Economic Behavior*, 49, 2004.

[24] M. Zinkevich, A. Greenwald, and M. Littman. Cyclic equilibria in Markov games. In Y. Weiss, B. Schölkopf, and J. Platt, editors, *Proceedings of the 18th NIPS*, pages 1641–1648. MIT Press, 2005.