General Revision Protocols in Best Response Algorithms for Potential Games [1]

Manuscript by Vitor Bosshard
Seminar Algorithmic Game Theory, ETH Zürich
22.04.2016

1 Introduction

It is well known that the best-response algorithm (BRA) converges to a (pure) Nash equilibrium of every potential game. However, there is one assumption that is made implicitly, namely that it is accurate or useful to model players as playing sequentially one after another in some sort of round-robin scheme. This paper deals with relaxing this constraint and seeing what happens.

As an example, consider a game between 2 players given by the following matrix:

1\2	a	b
a	0, 0	2, 2
b	3, 3	1, 1

This game is clearly an exact potential game, with potential function equal to the payoffs of each player. It follows that it converges under best response dynamics. However, if the players start at strategy profile (a, a) or (b, b) and move simultaneously, they will oscillate between those 2 non-equilibrium profiles indefinitely. If just one of the players is allowed to play alone, even if only very rarely, the algorithm converges to a Nash equilibrium immediately when such a move happens.

To analyze this behaviour, we need the concept of a revision protocol, which formalizes the possible turn orders of the players. Based on a property of the revision protocol called separability, it is possible to show exactly under which conditions a game converges to a Nash equilibrium.

2 Formal Definitions

2.1 Games and Best Responses

Definition 1. *A game is a triple* (N, A, u) *with*

- *A set of players* $N := \{1, \dots, n\}$
- *A set of action profiles* $A := \prod_k A_k$ *where* A_k *is a set of actions for player* k*.*

- *A set of utility functions* $u := \{u_1, \ldots, u_n\}$.

Definition 2. *The best response correspondence*

$$BR_k(x) := \underset{\alpha \in A_k}{\text{argmax}} \quad u_k(\alpha; x_{-k})$$

is the set of actions that maximizes the payoff for player k *under action profile* x.

The paper assumes that the best response correspondence is a function, i.e. that for every player and every action profile, there is a unique best response that player can play. This is not just a simplification; as it turns out this property is essential for certain proofs to hold.

The paper deals exclusively with a class of games called BR-potential games, which are a generalization of exact potential games. Every exact potential game is also a BR-potential game.

Definition 3. *A game is a BR-potential game if it admits a funcion* $F : A \mapsto \mathbb{R}$ *(called the potential) such that*

$$BR_k(x) := \underset{\alpha \in A_k}{\text{argmax}} \quad F(\alpha; x_{-k})$$

2.2 Revision Protocols

A revision protocol is one way of modelling a distributed best-response dynamic, where several players are reacting to their current situation in the game, without necessarily coordinating with other players.

Definition 4. *A revision protocol for a set* N *of players is a probability distribution* ρ *over all subsets* $K \subseteq N$*, such that every player is contained in at least one of the subsets with strictly positive probability.*

The sets with strictly positive probability are known as the support of ρ, $S(\rho)$ for short.

Note that this definition does not depend on the current strategy profile of the game or on any property of the game for that matter.

As a simple example, consider the situation where at each timestep, each player randomly chooses (with 50% probability) to either change their strategy to be a best response, or do nothing. This results in there being a small probability (namely 2^{-n}) of any given subset of players moving together.

Note that in this example, even though any group of players are capable of moving together, there is no explicit coordination going on at all, no central authority giving players permission to move. Arguably, this models many real world situations better than the more conventional best response algorithm.

2.3 Separability

Definition 5. *A revision protocol for a set of players N is separable if its support contains a sequence of sets $(K_1, \ldots, K_n)$ such that, for all i,*

$$K_i \setminus \bigcup_{j<i} K_j$$

is a singleton.

In other words, it is possible to take some player that appears in a singleton set, remove that player from all sets of the protocol, and end up with a separable revision protocol of the remaining $n-1$ players. This process could then continue by finding another singleton and removing the corresponding player, until no players remain.

Separability is a property independent of the probabilities of each set being chosen, only depending on the support of the distribution. Examples of separable supports include

- $\{\{1\}, \{2\}, \{3\}\}$,
- $\{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$,
- $\{\{1\}, \{2\}, \{1, 2, 3\}\}$,
- $\{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$.

3 Convergence to Nash Equilibria for Separable Revision Protocols

Now that the relevant definitions have been established, we move on to the first result of the paper.

Consider a modified best response algorithm where the set of players moving next are drawn according to a revision protocol, and then all selected players move to their best response simultaneously. If we denote the best response algorithm as BRA, from now on we must talk about $BRA(\rho)$ instead; the algorithm has become dependent on the revision protocol.

It turns out that separability is exactly the condition the protocol must fulfill for this process to always converge to a Nash equilibrium:

Theorem 1. *Let N be a set of players and ρ be a revision protocol over N. Algorithm $BRA(\rho)$ converges a.s. to a Nash equilibrium for all BR-potential games G over N if and only if ρ is separable.*

The proof is presented in 2 steps, first it is shown that the condition is sufficient and then that it is also necessary.

Proof of "$\Leftarrow$" (separability implies convergence to Nash equilibrium). We need to show this property to hold for an arbitrary game G. We construct a Markov chain from G and $BRA(\rho)$ as follows: Every action profile $x \in A$ is a node in the chain. Transition probabilities between nodes are given by the probability that the algorithm will move from one profile directly to the other.

Note that several subsets of players may add to the probability of one transition (intuitively, the Markov chain can be considered as a multigraph).

Now we define a set R consisting of all recurrent action profiles, that is those nodes in the chain that are visited infinitely often. The proof idea is to show that R is made exclusively of Nash equilibria.

We assume by contradiction that there exists a profile $x \in R$ that is not a Nash equilibrium, and we assume that x is the one with the highest potential among all possible choices for x.

From that profile, the sequence $(K_1, \ldots, K_n)$ of sets of players (as given in the definition of separability) has a positive probability of being chosen by the algorithm, generating the sequence $(X_1, \ldots, X_n)$ of action profiles. Note that some of the X_i may be equal to x.

Let i be the smallest index such that the action profile $X_i \neq x$. It follows that X_i is the best response of exactly one player to action profile x. This has to be the case, since by definition, all players who had the opportunity to deviate from x before the algorithm reached X_i, chose not to do so.

This implies that the potential strictly increases when going from x to X_i, because it is a unilateral deviation of only one player, and we are assuming that the best response is always unique. If the potential didn't strictly increase, it would imply that both x and X_i are best responses to x by that player.

Now we are almost done: if X_i is a Nash equilibrium, then x cannot be visited infinitely often (since it would lose probability mass to X_i over time). However, if X_i is not a Nash equilibrium, it contradicts our assumption that x has highest potential among all recurrent profiles.

□

Proof of "⇒" (convergence to Nash equilibrium implies separability). We need to show that any revision protocol ρ over N players which converges on every game is separable.

The proof is by induction, assuming the property to hold for every game of $n-1$ players. The induction base $n = 1$ is trivial.

First we show the result with the additional assumption that ρ contains a singleton that has positive probability. (This assumption is in fact true, as shown below).

We take some player k that occurs in a singleton, and remove them from all sets in ρ. This results in a revision protocol ρ' over $n-1$ players. ρ' converges for any game over $n-1$ players, because each such game G can be extended to a game G^+ over n players by adding a dummy player with only 1 action available. The sequence of states visited by $BRA(\rho)$ on G^+ is isomorphic to the sequence of states visited by $BRA(\rho')$ on G, and since $BRA(\rho)$ always converges, so must $BRA(\rho')$. By the IH, ρ' must therefore be separable, and by definition of separability, so must be ρ.

It remains to show that ρ must contain a singleton that has positive probability. For this purpose, it is enough to exhibit one specific game that cannot ever converge to a Nash equilibrium if ρ has no singletons.

This game, which we call G^*, is defined as follows:

Each of the n players has a set of actions $A_k := \{0, \ldots, p-1\}$, where p is any prime larger than N. The payoffs are symmetrical for every player and equal to the potential

$$F(x) := -(\sum_k x_k \mod p).$$

(Note that this way of constructing a game from a potential function works for any function F and always results in an exact potential game.)

From this it follows that from a state with potential $-h$ the best response for any player is to deviate by exactly h from their current strategy (using modular arithmetic, of course), that is

$$BR_k(x) := (x_k - h \mod p; x_{-k}).$$

If m players play simultaneously from a state with potential $-h$ this results in the next state having potential

$$\begin{aligned} &-h + mh \mod p \\ =&-(1-m)h \mod p, \end{aligned}$$

which is zero only when $m = 1$, because p isn't divisible by any number in the range $\{1, \ldots, n\}$.

Since h was chosen arbitrarily, this implies that if we start with any potential smaller than 0, we will never reach a Nash equilibrium, which all have potential 0 by construction. □

4 The Smoothed Best Response Algorithm

One of the typical weaknesses of any iterative optimization algorithm such as BRA is that they often get stuck in local optima, in our case Nash equilibria with less than maximal potential. One way to overcome this problem is to occasionally make deliberately suboptimal choices, which allows breaking free from any local optimum that might have otherwise trapped the algorithm.

This process is usually guided by a temperature parameter, which controls how willing the algorithm is to jump from a better to a worse state. If the temperature is decreased slowly over time, the algorithm converges eventually. The final value is often, but not always, the global optimum or a very good local optimum. This general approach is known as simulated annealing.

In the paper, a $SmoothBRA$ algorithm is specified which implements this idea. At each state x each player k, if chosen to act next, consults a probability distribution $Q_k(x)$ to determine which action to take, defined as follows:

$$\Pr[Q_k(x) := \alpha] = \frac{\exp(\theta u_k(\alpha; x_{-k}))}{\sum_{\beta \in A(k)} \exp(\theta u_k(\beta; x_{-k}))}$$

When the temperature $1/\theta$ is infinite, this rule chooses uniformly at random between all strategies. On the other extreme, as the temperature tends towards 0, the distribution tends towards only ever picking the best response.

The transition probabilities of the Markov chain induced by a game and a revision protocol can now be computed. The probability of an (x, y) edge, that is the probability that $BRA(\rho)$ will move directly from state x to state y, is given by the following formula:

$$P_{x,y} = \sum_{V \supseteq \mathrm{Diff}(x,y)} \rho(V) \prod_{k \in V} \frac{\exp(\theta u_k(\alpha; x_{-k}))}{\sum_{\beta \in A(k)} \exp(\theta u_k(\beta; x_{-k}))}$$

where $\mathrm{Diff}(x, y) := \{k \mid x_k \neq y_k\}$ is the set of players that must have moved if the state jumps from x to y. This can be rewritten as a first order approximation, that is, the highest order term is separated from lower order terms:

$$P_{x,y} = c_{x,y} \exp(\theta)^{q_{x,y}} + o(\exp(\theta)^{q_{x,y}})$$

Here, $q_{x,y}$ is called the *order* of $P_{x,y}$, and it will be relevant for later proofs. The order can be seen to be

$$q_{x,y} = \min_{V \supseteq \mathrm{Diff}(x,y) \cap S(\rho)} \left(\sum_{k \in V} \left(\max_{\alpha \in A_k} u_k(\alpha; x_{-k}) - u_k(y_k; x_{-k}) \right) \right),$$

where it should be recalled that $S(\rho)$ is the support of ρ. The order is always positive, and 0 exactly when it is possible to go from x to y using a set of players that all play a best response to x.

The concept of orders will turn out to be very relevant.

4.1 The Markov Chain Tree Theorem and Stochastic Stability

The paper makes use of a classic result established in [2] which allows to determine the stationary distribution π of any Markov chain asymptotically.

Theorem 2 (Markov Chain Tree Theorem)**.** *Let T_x be the set of spanning trees in the transition graph with root in x. The stationary probability π_x is proportional to the sum of the weights of all such trees, i.e.*

$$\pi_x \propto \sum_{T \in T_x} \prod_{(y,z) \in T} P_{y,z}.$$

The type of Markov chain we are working with has the property that for any action profile x, there is a positive probability that it will be visited infinitely often.

However, as we observe the asymptotic behaviour of the chain as $\theta \to \infty$, some profiles will stabilize at a positive probability (while others go to 0). These profiles are called stochastically stable under ρ.

The stochastically stable profiles can be characterized by the following lemma, which follows directly from the Markov chain tree theorem:

Lemma 1 (Stochastic stability characterization)**.** *Let*

$$q_x := \min_{T \in T_x} \sum_{(y,z) \in T} q_{yz}.$$

be the order of the minimal in-tree of state x.
x is stochastically stable if and only if q_x is minimal, i.e.

$$q_x \in \min_{y \in A} q_y.$$

4.2 Convergence to Nash Equilibria for Smoothed BRA

As mentioned above, for any fixed temperature $1/\theta$, there must remain a small residue of probability mass even in states with very low potential. The low potential state loses almost all its mass to other states at each timestep, but states with high potential lose a small fraction of their mass to that state as well. At some point, those 2 flows will be balanced, resulting in a very low (but not 0!) stable probability.

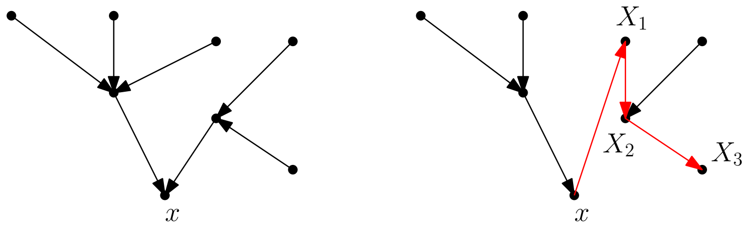

Figure 1: Example of rerouting a tree to a different root, as used in the proof of Theorem 3.

Intuitively, running the algorithm with high temperature ensures fast convergence to the stationary distribution because low potential states will bleed mass very fast, but the final distribution will have a larger amount of mass in non-equilibrium states, because the equilibria are bleeding a lot as well. Conversely, running with a low temperature may take a longer time to converge, but results in a better stationary distribution.

If we are going to talk about convergence for the $SmoothBRA$ algorithm, we can only make a claim as to the set of stochastically stable states. As it turns out, $SmoothBRA$ always converges to a Nash equilibrium in that weaker sense.

Theorem 3 (Convergence to Nash equilibrium)**.** *Let G be a BR-potential game, and ρ a separable revision protocol. An acion profile x of G cannot be stochastically stable under ρ if it is not a Nash equilibrium.*

Proof. Suppose we run the non-smoothed version of the algorithm, $BRA(\rho)$, starting at a state x that is not a Nash equilibrium. Then, it follows from Theorem 1 that there is a finite sequence of action profiles $(X_1, X_2, \ldots, X_H)$ leading to a Nash equilibrium X_H. This sequence creates a path of order 0, since it only consists of best responses by sets of players.

Let T_x^* be the tree with minimal order rooted in x. We construct a tree rooted in X_H by adding the path defined above and removing all other edges outgoing from vertices in the path. See Figure 1 for an example.

The added edges all have order 0, while at least one removed edge has order strictly greater than 0, namely the one leading out of X_H. Therefore, this new tree has order strictly smaller than T_x^*, so by Lemma 1 x is not stochastically stable.

□

4.3 Nonconvergence to Optimal Nash Equilibria

We might hope that if we run the smoothed version of the BRA algoritm, it always converges to the optimal Nash equilibrium. Unfortunately, this is not true for all games and revision protocols. Here we provide an example that illustrates what can go wrong:

We have a 3-player game where every player has 2 actions. The revision protocol has the support

$$S(\rho) = \{\{1\}, \{2\}, \{3\}, \{1, 2, 3\}\}.$$

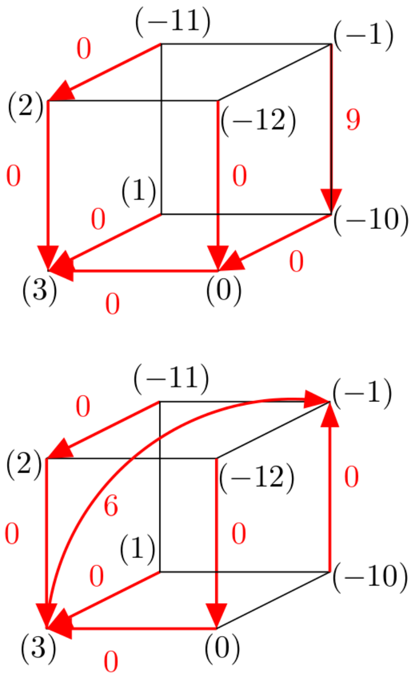

Figure 2: Example of a 3-player game that doesn't converge to an optimal NE. The payoffs are given in parenthesis. The minimal in-tree rooted in $(0, 0, 0)$ has order 9, while the one rooted in $(1, 1, 1)$ has order 6. Figure taken from the paper.

As in the previous example, the payoffs are symmetrical for each player and equal to the potential. They are given in parentheses in Figure 2, where each action profile x is represented as a vertex of the 3-cube.

The Nash equilibria are the antipodals $(0, 0, 0)$ and $(1, 1, 1)$. However, the minimal tree rooted in $(0, 0, 0)$ has order 9, while the one rooted in $(1, 1, 1)$ has order 6. Therefore, only $(1, 1, 1)$ is stochastically stable, even though it doesn't have optimal potential.

What happens here is that the non-local nature of the diagonal jump completely ignores the bad neighborhood around $(1, 1, 1)$. If simultaneous play was not possible, a huge penalty would have to be paid to reach any of $(1, 1, 1)$'s three neighbors.

4.4 Special Cases

In some special cases it is possible to prove convergence to the optimal Nash equilibrium by the smoothed algorithm. Two such cases are given in the paper. Here we just briefly mention them without proof.

Theorem 4 (Convergence to optimal Nash equilibria for asynchronous revisions)**.** *Let G be an exact potential game. Under a revision protocol ρ with support*

$$S(\rho) = \{\{1\}, \{2\}, \ldots, \{n\}\},$$

containing only singletons, the only stochastically stable states are the optimal Nash equilibria.

Theorem 5 (Convergence to optimal Nash equilibria with 2 players)**.** *Let G be an exact potential game between 2 players. Under a revision protocol ρ with support*

$$S(\rho) = \{\{1\}, \{2\}, \{1, 2\}\},$$

the only stochastically stable states are the optimal Nash equilibria.

References

[1] Pierre Coucheney, Stéphane Durand, Bruno Gaujal, and Corinne Touati. General revision protocols in best response algorithms for potential games. In *Netwok Games, Control and OPtimization (NetGCoop)*, 2014.

[2] Frank Thomson Leighton and Ronald L Rivest. *The Markov chain tree theorem.* Massachusetts Institute of Technology, Laboratory for Computer Science, 1983.