Convergence to a Convention in a 2x2 Coordination Game with Adaptive Learning

Ethan Holdahl
University of Oregon
April 5, 2022

Coordination Game G

		Player B	
		0	1
Player A	0	$a_{00}, b_{00}{}^*$	a_{01}, b_{01}
	1	a_{10}, b_{10}	$a_{11}, b_{11}{}^*$

Lemma 1 will be used in the proof of Theorem 1.

Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. Denote the players A and B with strategies $\{0, 1\}$ as depicted above where (0,0) and (1,1) constitute pure strategy Nash Equilibria. Define $h^t_{-i} = (x^{t-m}_{-i}, .., x^{t-1}_{-i})$ as the most recent m records of all players except player i at time t. Define R^t_i as the set of s records sampled by player i in period t from h^t_{-i}. Define BR^t_i as player i's best response to R^t_i.

Lemma 1. Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ. If at any period n $\exists$ a strategy profile $x^* = (x^*_1, x^*_2)$ that constitutes a Nash Equilibrium where $x^*_1 \in \{BR^n_1\}$ and $x^*_2 \in \{BR^n_2\}$ for players 1 and 2 then there exists a positive probability that each player i plays x^*_i as a best response for every period $e \geq n$.

Proof of Lemma 1. Assume at period n $\exists$ an action $x^* \in BR^n_A \cap BR^n_B$. I must show that there exists a positive probability that each player plays x^* as a best reply for every period $e \geq n$.

I use proof by induction.

Base Step: I show that if there exists an action x^* that is a best response for both players in period n then there exists a positive probability that both players play action x^* in period n.

Clearly, if x^* is a best response for both players then there exists a positive probability that both players play action x^* in period n.

Inductive Step: I show that $\forall e \geq n$ if action x^* was played as a best response for both players in period e then there exists a positive probability that action x^* is played as a best response in period $e+1$.

Assume action x^* was played as a best response for both players in period e. I must show that there exists a positive probability that action x^* is played as a best response in period $e+1$.

Note that since $x_i^t \in \{0,1\} \quad \forall t$, the proportion of times that player $j \neq i$ played 1 in R_i^t, the set of s records in period t, is simply $\sum_{r \in R_i^t} \frac{r}{s}$.

For example, if $R_i^t = (0,1,1,1,0,0,1,1)$ then then proportion of times player $j \neq i$ played 1 in player i's sample, $R_i^t = \sum_{r \in R_i^t} \frac{r}{s} = \frac{5}{8}$

Without loss of generality assume that $x^* = 1$. That means action 1 is a best response for each player $i \in \{A, B\}$ to R_i^e, the set of s records sampled by player i in period e.

Let $\alpha_i \in (0,1)$ be the smallest probability that Player $j \neq i$ plays action 1 such that Player i's best response is playing action 1.
That means that for $i \in A, B$:

$$(1) \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

Now consider the set of s records sampled by player i in $e+1$: R_i^{e+1}.

In period $e+1$ each player i samples s records from h_{-i}^{e+1}, the most recent m records of all players except player i at time t. Note that $|h_{-i}^e \cap (h_{-i}^{e+1})'| = 1$, that is to say that there is only 1 record in h_{-i}^e that is not in h_{-i}^{e+1}. Since $R_i^e \subseteq h_{-i}^e$ I know that $|R_i^e \cap (h_{-i}^{e+1})'| \leq 1$. That means there is at most 1 record in player i's sample in period e that is not able to be sampled in period $e+1$. This means that there is a positive probability that in period $e+1$, each player i samples $s-1$ records from the set $h_{-i}^{e+1} \cap R_i^e$ and the most recent record, x_{-i}^e. Assume both players samples in period $e+1$ fit this criteria and define $c_{-i}^e = R_i^e \backslash R_i^{e+1}$, the record that was in the sample in period e but not in period $e+1$ for player i.

So, $c_{-i}^e = R_i^e \backslash R_i^{e+1}$ and $x_{-i}^e = R_i^{e+1} \backslash R_i^e$. Consequently, I know that for each player i:

$$\frac{c_{-i}^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} = \frac{x_{-i}^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

However, I know that $x_{-i}^e = 1$ and $c_{-i}^e \in \{0, 1\}$. So, $x_{-i}^e \geq c_{-i}^e$.
Adding to both sides I get:

$$\frac{c_{-i}^e}{s} + \frac{x_{-i}^e}{s} + \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \frac{x_{-i}^e}{s} + \frac{c_{-i}^e}{s} + \sum_{r \in R_i^e} \frac{r}{s}$$

So

$$(2) \sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s}$$

Using (1) and (2) by transitivity I get:

$$\sum_{r \in R_i^{e+1}} \frac{r}{s} \geq \sum_{r \in R_i^e} \frac{r}{s} \geq \alpha_i$$

So for each player $i \in \{A, B\}$ action 1 is a best response to the sample R_i^{e+1}. Since action 1 is a best response for both players in period $e + 1$ there exists a positive probability that each plays action 1 in period $e + 1$.

Since both the base case and the inductive step has been shown, by mathematical induction I have proven that if both players play x^* as a best response in period n, then there exists a positive probability that both players play x^* as a best reply for every period $e \geq n$. $\square$

Theorem 1. Let G be a 2x2 coordination game, and let $P^{m,s,\epsilon}$ be adaptive learning with memory m, sample size s, and error rate ϵ.

If $s < m$ then from any initial state, the unperturbed process $P^{m,s,0}$ converges with probability one to a convention and locks in.

Proof of Theorem 1. Define G as a 2x2 coordination game with adaptive learning where the possible actions for both players A and B are $\{0, 1\}$. Let memory $m \in \mathbb{N}$, sample size $s \in \mathbb{N}$ such that $s < m$, error rate $\epsilon = 0$ and let $h^t = (x^{t-m+1}, ..., x^t)$, be an arbitrary state at the end of period t. Let $\alpha \in (0, 1)$ be the smallest probability that Player B plays action 1 such that Player A's best response is playing action 1. Likewise, Let $\beta \in (0, 1)$ be the smallest probability that Player A plays action 1 such that Player B's best response is playing action 1.

There exists a positive probability that both players sample the most recent set of s records: $\{x^{t-s+1}, ..., x^t\}$ in period $t + 1$. Assume this is the case.

In period $t + 1$ the two players either
1) Share a best reply
or
2) Do not share a best reply
I will show a convention can be reached with positive probability in both cases.
Case 1: Both players share a best reply, x^*, in period $t + 1$. In this case I can apply Lemma 1 which shows that there exists a positive probability that both players play action x^* as a best reply for each period $e \geq t + 1$. If this happens then after period $e = t + m$ the entire memory is filled with both players playing action x^*. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing x^* both would continue to play x^* as a best response for every period thereafter. So, I have shown that there exists a positive probability that a convention can be reached and locked into with positive probability in Case 1.

Case 2: Assume the players do not share a best reply to the most recent set of s records. Since $\epsilon = 0$ they play different actions as best replies in period $t + 1$. Without loss of generality assume that in period $t + 1$ player A played action 1 and player B played action 0. This means that:

$$(3) \sum_{r=t-s+1}^{t} \frac{x_B^r}{s} > \alpha$$

and

$$(4) \sum_{r=t-s+1}^{t} \frac{x_A^r}{s} < \beta$$

Note: this is a strict inequality since the players do not share a best reply in period $t+1$ after sampling the set of records: $(x^{t-s+1}, ..., x^t)$.

Defining k and j

Assume for the time being that player B continues to play action 0 for every period after period $t+1$. We know, since this is a coordination game, that if player A samples the most recent s actions in every period that there will exist a period, let's call the first one period $t+1+k$, where player A will have action 0 as a best response.

Thus, k is defined to be the smallest integer such that

$$\sum_{r=t-s+1+k}^{t+k} \frac{x_B^r}{s} \leq \alpha$$

Note that we assume that the record of x_B^r is 0 for $r > t+1$. So, the sum of the records x_B^r where $r > t+1$ is equal to 0 and drops out of this best response calculation.
So, the above equation can be simplified to:

$$(5) \sum_{r=t-s+1+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$$

Since α is positive we know the inequality holds when $k = s$, as that would force the right side of the equation to zero since the record of $x_B^{t+1} = 0$. Additionally, (3) implies that it does not hold when $k = 0$. Thus, it is clear that for all histories and all α, $k \in \{1, ..., s\}$.

Likewise, let's assume for the time being that player A continues to play action 1 for every period after period $t+1$. We know, since this is a coordination game, that if player B samples the most recent s actions in every period that there will exist a period, let's call the first one period $t+1+j$, where player B will have action 1 as a best response.

Thus, j is defined to be the smallest integer such that

$$\sum_{r=t-s+1+j}^{t+j} \frac{x_A^r}{s} \geq \beta$$

Note that we assume that the record of x_A^r is 1 for $r > t+1$. So, the sum of the records $\frac{x_A^r}{s}$ where $r > t+1$ is equal to $\frac{j-1}{s}$.
So, the above equation can be simplified to:

$$(6)\frac{j-1}{s} + \sum_{r=t-s+1+j}^{t+1} \frac{x_A^r}{s} \geq \beta$$

Since β is less than 1 we know the inequality holds when $j = s$ since the right side of the equation equals 1 since the record of $x_A^{t+1} = 1$. Also, (4) tells us that it does not hold when $j = 0$. Thus, it is clear that for all histories and all β, $j \in \{1, ..., s\}$.

Note: allowing j = 0 is not problematic in (6) because we know $x_A^{t+1} = 1$, which will cancel out with $\frac{j-1}{s}$ leaving the expression to be equal to that in (4).

Now I will prove that for all periods after $t+1$ and before $t+1+k$ there is a positive probability that player A's best response is to play action 1. Likewise, I will prove that for all periods after $t+1$ and before $t+1+j$ there is a positive probability that player B's best response is to play action 0.

Proving best responses in periods $t+1+e$ $\forall e$ such that $0 < e < k, j$

If $k = 1$ then there is nothing to prove with respect to player A as there is no integer between 0 and 1. So, I will prove that 1 can be a best response for player A in periods $t+1+e$ $\forall 0 < e < k$ where $k > 1$.

Likewise, if $j = 1$ then there is nothing to prove with respect to player B as there is no integer between 0 and 1. So, I will prove that 0 can be a best response for player B in periods $t+1+e$ $\forall 0 < e < j$ where $j > 1$.

Assume that in each period $t+1+e$ that players samples the most recent set of s records. In this case, player A has action 1 as a unique best response if:

$$(7) \quad \sum_{r=t-s+1+e}^{t+e} \frac{x_B^r}{s} > \alpha$$

and player B has action 0 as a unique best response if:

$$(8) \qquad \sum_{r=t-s+1+e}^{t+e} \frac{x_A^r}{s} < \beta$$

Note that $j, k \leq s$ so $e \leq s-1$, and the restriction that $k, j > 1$ implicitly restricts s to $s > 1$. Because we are only concerned with cases where $1 \leq e \leq s-1$, and we know that $s > 1$, we can dissect the summation into 2 parts without consequence:

$$(9) \qquad \sum_{r=t-s+1+e}^{t+e} \frac{x_i^r}{s} = \sum_{r=t-s+1+e}^{t} \frac{x_i^r}{s} + \sum_{r=t+1}^{t+e} \frac{x_i^r}{s}$$

So combining equations (7) and (9) where $i = B$, player A has action 1 as a unique best response if:

$$(10) \quad \sum_{r=t-s+1+e}^{t} \frac{x_B^r}{s} + \sum_{r=t+1}^{t+e} \frac{x_B^r}{s} > \alpha$$

and combining equations (8) and (9) where $i = A$, player B has action 0 as a unique best response if:

$$(11) \quad \sum_{r=t-s+1+e}^{t} \frac{x_A^r}{s} + \sum_{r=t+1}^{t+e} \frac{x_A^r}{s} < \beta$$

Consider $1 \leq e < k$. Since k is the smallest integer such that

$$(5) \quad \sum_{r=t-s+1+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$$

it follows that

$$\forall e < k, \sum_{r=t-s+1+e}^{t+1} \frac{x_B^r}{s} > \alpha$$

Further, since $x_B^{t+1} = 0$, and $e \leq s-1$, the expression

$$\sum_{r=t-s+1+e}^{t} \frac{x_B^r}{s} = \sum_{r=t-s+1+e}^{t+1} \frac{x_B^r}{s} > \alpha$$

Since

$$x_B^r \in [0,1] \ \ \forall r, \sum_{r=t+1}^{t+e} \frac{x_B^r}{s} \in [0, \tfrac{e}{s}]$$

So, because

$$\sum_{r=t+1}^{t+e} \frac{x_B^r}{s} \geq 0 \ \text{ and } \sum_{r=t-s+1+e}^{t} \frac{x_B^r}{s} > \alpha$$

I get the condition (10):

$$\sum_{r=t-s+1+e}^{t} \frac{x_B^r}{s} + \sum_{r=t+1}^{t+e} \frac{x_B^r}{s} > \alpha$$

This condition means that for all integers e such that $1 \leq e < k$ when sampling the most recent s records in period $t+1+e$ that action 1 is a unique best response for player A.

The proof for player B playing 0 as a best response is similar to the one above:

Now consider $0 \leq e < j$. Since j is the smallest integer such that

$$(6)\frac{j-1}{s} + \sum_{r=t-s+1+j}^{t+1} \frac{x_A^r}{s} \geq \beta$$

it follows that

$$\forall e < j, \frac{e-1}{s} + \sum_{r=t-s+1+e}^{t+1} \frac{x_A^r}{s} < \beta$$

Further since $x_A^{t+1} = 1$, and $e \leq s-1$, the expression

$$\frac{e}{s} + \sum_{r=t-s+1+e}^{t} \frac{x_A^r}{s} = \frac{e-1}{s} + \sum_{r=t-s+1+e}^{t+1} \frac{x_A^r}{s} < \beta$$

Since

$$x_A^r \in [0,1] \ \ \forall r, \sum_{r=t+1}^{t+e} \frac{x_A^r}{s} \in [0, \tfrac{e}{s}]$$

So, because

$$\sum_{r=t+1}^{t+e} \frac{x_A^r}{s} \leq \frac{e}{s} \text{ and } \frac{e}{s} + \sum_{r=t-s+1+e}^{t} \frac{x_A^r}{s} < \beta$$

I get the condition (11):

$$\sum_{r=t-s+2+e}^{t} \frac{x_A^r}{s} + \sum_{r=t+1}^{t+e} \frac{x_A^r}{s} < \beta$$

This condition means that for all integers e such that $0 \leq e < j$ when sampling the most recent s records in period $t+2+e$ that action 0 is a unique best response for player B.

There exists a positive probability that in each period $t+1+e$ both players can sample the most recent s records. I have just shown that when sampling the most recent s records in periods $0 < e < k$, player A best responds with action 1. When sampling the most recent s records in periods $0 < e < j$, player B best responds with action 0. Assume both players do sample the most recent s records in those periods. Since $\epsilon = 0$ both players play their best response in those periods.

Proving coordination under different scenarios

I will now consider the three scenarios: $j < k$, $k < j$, and $j = k$.

First, $j < k$.

For period $t+1+j$ both players have a positive probability of sampling the most recent s records: $(x^{t-s+1+j}, ..., x^{t+j})$. Player B has a best response of action 1 in period $t+1+j$ if:

$$(12) \quad \sum_{r=t-s+1+j}^{t} \frac{x_A^r}{s} + \sum_{r=t+1}^{t+j} \frac{x_A^r}{s} \geq \beta$$

Since $j < k$, I have already shown that there is a positive probability that player A has a unique best response of playing action 1 for all periods between $t+1$ and $t+1+j$ inclusive. So, $\sum_{r=t+1}^{t+j} \frac{x_A^r}{s} = \frac{j}{s}$.

Which means

$$\sum_{r=t-s+1+j}^{t} \frac{x_A^r}{s} + \sum_{r=t+1}^{t+j} \frac{x_A^r}{s} = \frac{j}{s} + \sum_{r=t-s+1+j}^{t} \frac{x_A^r}{s}$$

Since we know $x_A^{t+1} = 1$, we know

$$\frac{j}{s} + \sum_{r=t-s+1+j}^{t} \frac{x_A^r}{s} = \frac{j-1}{s} + \sum_{r=t-s+1+j}^{t+1} \frac{x_A^r}{s}$$

Using (6), we know:

$$\frac{j-1}{s} + \sum_{r=t-s+1+j}^{t+1} \frac{x_A^r}{s} \geq \beta$$

which means that (12) holds and action 1 is a best response for Player B in period $t+1+j$.

Since both players can, with positive probability, sample the most recent s records and have the same action as a best response in period $t+j+1$, I know by Case 1 a convention can be reached and locked into with positive probability when $j < k$.

Second, I consider $k < j$.

For period $t+1+k$ both players have a positive probability of sampling the most recent s records: $(x^{t-s+1+k}, ..., x^{t+k})$. Player A has a best response of action 0 in period $t+1+k$ if:

$$(13) \sum_{r=t-s+1+k}^{t} \frac{x_B^r}{s} + \sum_{r=t+1}^{t+k} \frac{x_B^r}{s} \leq \alpha$$

Since $k < j$, I have already shown that player B has a unique best response of playing action 0 for all periods between $t+1$ and $t+1+k$ inclusive. So $\sum\limits_{r=t+1}^{t+k} \frac{x_B^r}{s} = 0$.
Which means

$$\sum_{r=t-s+1+k}^{t} \frac{x_B^r}{s} + \sum_{r=t+1}^{t+k} \frac{x_B^r}{s} = \sum_{r=t-s+1+k}^{t} \frac{x_B^r}{s}$$

Since we know $x_B^{t+1} = 0$, we know

$$\sum_{r=t-s+1+k}^{t} \frac{x_B^r}{s} = \sum_{r=t-s+1+k}^{t+1} \frac{x_B^r}{s}$$

Using (5), we know:

$$\sum_{r=t-s+1+k}^{t+1} \frac{x_B^r}{s} \leq \alpha$$

which means that (13) holds and action 0 is a best response for Player A in period $t+1+k$.

Since both players can, with positive probability, sample the most recent s records and have the same action as a best response in period $t+k+1$, I know by Case 1 a convention can be reached and locked into with positive probability when $k < j$.

Third, I consider $k = j$.

In period $t+1+k$ player B can, with positive probability, sample the most recent set of s records. Since $j = k$ I have already shown that playing action 1 is a best response for player B in this scenario.

Since $m > s$ and both m and s are integers, I know that $m \geq s+1$. So in period $t+1+k$ the records player A can sample from includes the most recent $s+1$ records: $(x^{t+k-s}, ..., x^{t+k})$. So in period $t+1+k$ player A can, with positive probability, sample the set of s records: $(x^{t+k-s}, ..., x^{t+k-1})$. Note that these are the same records sampled in the previous period, $t+k$, by player A which gave the unique best response of playing action 1.

So, there exists a positive probability that both players share a best response in period $t+1+k$. As consequence, I can apply Lemma 1 here which shows that there exists a positive probability that both players play action 1 as a best reply for each period $e \geq t+1+k$. If this happens, then after period $t+k+m$ the entire memory is filled with both players playing action 1. Since $\epsilon = 0$ and since both players could then only sample records of the other player playing action 1 both would continue to play 1 as a best response for every period thereafter. So I have shown that there exists a positive probability that a convention can be reached and locked into when $j = k$.

Thus, I have exhausted all three scenarios: $j < k$, $k < j$, and $j = k$ and shown that a convention can be reached with a positive probability in Case 2.

Since I have shown that a convention can be reached with positive probability in both Case 1 and Case 2 I have proven that from any initial state when $s < m$ and $\epsilon = 0$ a convention can be reached with positive probability and lock in. Since a convention can be reached from any arbitrary state and since conventions are absorbing states we know that as $T \to \infty$ that h^T, the state at time T converges with probability one to a convention in and locks in. ■