How to efficiently defeat a strategy of bounded rationality

Guillaume LACÔTE[1] and Grégory THURIN[2]

Abstract : In a given finite two-player zero-sum game if player one is sufficiently bounded in rationality it might be possible for player two to beat him, *i.e.* to ensure that he never gets more than his maxmin payoff in pure strategies at nearly all stages of the repeated game. The issue in this case is first to determine the relative minimum bound in complexity required for player two to defeat player one; it is then to determine how well player two performs, that is at how many stages he fails to beat player one and after which stage (if any) he can be certain to defeat him at each stage. Elaborating on recent results this paper addresses these issues in the case where bounded rationality is alternatively specified by means of finite state automata or bounded recall strategies.

Keywords : Repeated games, bounded rationality, finite automata, bounded recall strategies.

1 Introduction

The soundness of many results of repeated games theory as a modelization of economic behaviors is moderated by their strong hypothesis that agents are individually rational and infinitely intelligent. Thus numerous recent works have focused on repeated games played by boundly rational agents (see [Rub98], [Kal93] or [Au81] for surveys).

In particular given a two-players game if player one is sufficiently bounded in rationality it might become possible for player two to *defeat* him (*i.e.* play a best-reply action to his action) at each stage of the associated repeated game. A key question is then to determine such a bound explicitely. This requires defining a concept of bounded rationality first.

The reasoning complexity of a strategy can be modelized through the Theory of computational models (see [HU79] for an introduction). Among those models the class of *Finite State Automata* has been shown (for instance by [Au81] or [Rub88]) to be particularly adequate to modelize rationality in repeated games. Roughly speaking a finite state automaton is a deterministic machine with a finite number of states; each state determines the action to be played, and after each stage the machine may change its state according to the last action played by the adverse player. Such an automaton induces a strategy [3], and thus the rationality of a given strategy can be defined by the minimum number of states of an automaton that can implement it (if such an automaton exists).

[1]ENSAE. 3 av. P. Larousse, F-92240 Malakoff FRANCE. Mail: `Games@glacote.com`
[2]UNIV. PARIS 1. P. du M^{al} de Lattre de Tassigny, F-75775 Paris FRANCE. Mail: `Games@thurin.com`
[3]More precisely a class of equivalence of strategies; refer to [Ney98] for details.

Under this context E. Ben-Porath considers an infinitely repeated two-players zero-sum game in which player one is restricted to strategies that can be implemented by an automaton of a given number $n \geq 1$ of states. The question is then to determine the minimum number of states of an automaton for player two whose induced strategy would beat *any* such strategy of player one without first knowing it. In [Por93] it is actually shown that there exists an automaton with in the order of n^{Cn+1} states which beats any strategy for player one implemented by an automaton of at most n states at every stage except at a finite number of them (where C is a constant which depends solely on the sets of actions of the one-shot game). The proof of the existence of this automaton is based on the fact that given a known automaton with n states for player one a best-reply automaton for player two, which plays a best-reply action at each stage, needs not be larger than n. The construction thus mainly consists in exhaustively concatenating all possible best-reply automata; the result is then derived from the fact that there are in the order of n^{Cn} such automata (see [Por93] for details).

On the other hand A. Neymann and D. Okada consider an infinitely repeated two-players game in which player one is restricted to a given set of (pure) strategies, be they implemented by an automaton or not. Player two knows this set and is supposed to be unbounded in complexity, *i.e.* he is allowed to play any (pure) strategy. It is legitimate under this context to hypothesize that due to its unbounded rationality player two will beat player one. The main issue turns out to be at how many stages (depending on the size of E) player one can actually avoid being beaten. In [NO00] it is proven that there exists a strategy for player two that beats any strategy of a given set E at any stage except possibly at $\lceil \log_2 |E| \rceil$ stages. Such a strategy is explicitely built based on the following observation : at any stage of the repeated game, among all actions that player one may play at next stage one of them is played more often than the others. Thus by playing a best-reply to this action, player two ensures that either player one will play it (and will thus get beaten), or that at most twice less strategies remain compatible with the history of played action than at the beginning of the stage. The second case may eventually lead player two to fully identify player one's strategy, and thus beat him at each subsequent stage. Note however that this winning strategy for player two is expressed in [NO00] in a functionnal form rather than through an automaton.

Elaborating on these results the purpose of this paper is to answer the following questions :

(i) assuming that player one is of bounded rationality, what is the minimum rationality required for player two to "*beat*" him ?

(ii) what is in this scenario the number of possible failures, *i.e.* the total number of stages at which player one will not get beaten ?

(iii) after which date (if any) can player two be certain to beat player one at each subsequent stage ?

More precisely this paper provides a (partial) unification of the aforementionned results in the following way : on the one hand it exhibits a winning automaton for player two which is at most twice larger than the one proposed by [Por93] and which wins logarithmically quicker. And on the other hand we built an *automaton* that implements a similar strategy as [NO00], although this requires a stronger hypothesis (hence the term *partial* unification). In addition these results are extended to the context of *Bounded Recall Strategies* : this article answers questions (i), (ii) and (iii) in the four possible cases where the rationality of each player is expressed either in terms of finite state automata or in terms of bounded recall.

The rest of this paper is organized as follows : the next section briefly recalls the notions of repeated games and finite automata. In section 3 we define the notion of cyclicity and state the main result. It is proven in section 4, and in section 5 it is extended to the context of finite automata and bounded recall strategies. A real-world example is given in section 6, and section 7 suggests possible extensions and future work.

2 Repeated games and finite automata

2.1 Repeated games

Let $G = (A^1, A^2, r)$ be a finite two-player game in strategic form, where A^i is the finite set of *actions* of player i and $r : A \to \mathbb{R}^2$ the *payoff* function, where $A = A^1 \times A^2$. For any player $i \in \{1,2\}$ we denote $'' - i'' = 3 - i$ the other player, and $u^i(G) = \max_{a^i \in A^i} \min_{a^{-i} \in A^{-i}} r^i(a^i, a^{-i})$ the *maxmin* in pure strategies; this is also the best payoff player i can guarantee in G with a pure strategy. For any given adverse action $a^{-i} \in A^{-i}$ let $\mathcal{B}r^i\left(a^{-i}\right)$ be an arbitrarily selected pure best reply to a^{-i}.

Let us define for $T \in \mathbb{N}^*$ the *finitely repeated game* G_T induced by the repetition of G T times, and the *infinitely repeated game* G_∞. First for $t \in \mathbb{N}^*$ let $H_t = A^{t-1}$ be the set of all possible *histories* at stage t, with the convention $A^0 = \{\oslash\}$ where $\oslash$ denotes the empty history. We call $H(T) = \bigcup_{t \in [\![1,T]\!]} A^{t-1}$ the set of histories *up to* stage T, and $H_\infty = \bigcup_{t \in \mathbb{N}^*} A^{t-1}$ the set of all histories. A *strategy* for player i in G_T is any $\sigma^i : H(T) \to A^i$; we note $S^i_T = \left(A^i\right)^{H(T)}$ the set of all such (pure) strategies. Similarly a strategy for player i in G_∞ is any $\sigma^i : H_\infty \to A^i$ and let $S^i = \left(A^i\right)^{H_\infty}$ be the set of all such strategies.

A profile of strategy $\sigma = (\sigma^1, \sigma^2)$ in G_T (respectively in G_∞) induces a *play* $\omega(\sigma)$ in H_T (respectively in H_∞) defined recursively in the following way : first let $\omega(\sigma)_1 = (\sigma^1_0(\oslash), \sigma^2_0(\oslash))$, and then set

$$\omega(\sigma)_t = \left(\sigma^1_{t-1}\left(\omega(\sigma)_{t-1}\right), \sigma^2_{t-1}\left(\omega(\sigma)_{t-1}\right)\right)$$

for any $t \in [\![2,T]\!]$ (respectively for any $t \geq 2$). Two strategies σ^i and τ^i for player i are *equivalent* if $\forall \sigma^{-i}\, \omega(\sigma^i, \sigma^{-i}) = \omega(\tau^i, \sigma^{-i})$. Finally let $r_T(\sigma) = \frac{1}{T}\sum_{t=1}^{T} r\left(\omega(\sigma)_t\right)$ be the *payoff* in G^T, and $r_\infty(\sigma) = \lim_{T\to+\infty} \frac{1}{T}\sum_{t=1}^{T} r\left(\omega(\sigma)_t\right)$ the payoff in G_∞. [4] This completes the definition of $G_T = (S_T, r_T)$ and $G_\infty = (S_\infty, r_\infty)$. A detailed study of repeated games can be found in [Sor86].

In the rest of this paper we will need the following additional definitions. For any play $\omega = (a_1, \ldots, a_t) \in H_t$ and any $l \leq t$ let $\omega_{|l} = (a_1, \ldots, a_l)$ be ω *truncated* at sage l, and note $\omega' \,[\![\!< \omega$ whenever ω' is a truncature of ω. In a similar way any $\alpha \in H_\infty$ is said to be *in tail* of ω, which is denoted by $\alpha \lhd\!]\!] \,\omega$, whenever $\exists \beta \in H_\infty \,/\, \beta\alpha = \omega$. Besides for $n \in \mathbb{N}$ and $h^i \in (A^i)^n$ let

$$\underline{h^i} = \left(\left(h^i_1, \mathcal{B}r^{-i}\left(h^i_1\right)\right), \ldots, \left(h^i_n, \mathcal{B}r^{-i}\left(h^i_n\right)\right)\right)$$

be the *best-reply path* of h^i.

2.2 Finite automata

Given a one-shot two-player game G we call finite state automaton (or *automaton* for short) for player i any $M^i = [Q^i, q^i_*, f^i, g^i]$ where :

- Q^i is a finite set called the set of *states* of the automaton
- $q^i_* \in Q^i$ is its *initial state*
- $f^i : (Q^i \times A^{-i} \to Q^i)$ is its *transition function*
- $g^i : (Q^i \to A^i)$ is its *action function*

The idea behind the model is that at each step the current state of the automaton determines its next action (through g^i); and after each step it might change its state according to the actions of other player (through f^i). Note that the current action does *not* depend directly on the past actions of the other player though. Formally given a strategy σ^{-i} for the other player an automaton M^i *induces* a strategy $\sigma(M^i)$ in the repeated game G_T (or G_∞) in the following way : [5]

- at first, M^i is in state $q^i_0 = q^i_*$ and plays actions $a^i_0 = g^i(q^i_0)$, while the other player plays action $a^{-i}_0 = \sigma^{-i}(\oslash)$
- the action a^{-i}_0 is then published to M^i, which changes its state to $q^i_1 = f^i(q^i_0, a^{-i}_0)$
- player i plays $a^i_1 = g^i(q^i_1)$ while the other player plays $a^{-i}_1 = \sigma^{-i}\left(\left(a^i_0, a^{-i}_0\right)\right)$
- and so on ...

[4] $r_\infty(\sigma)$ may not always exist in the general case. However along this article all considered strategies will be such that it is the case.

[5] This does *not* define any action after an history which could not have been played by the automaton, although a strategy should do it. Thus this defines only an *equivalence class* of strategies; but since in this paper automata will only face histories they have generated themselves we shall identify $\sigma(M^i)$ to a plain strategy.

Note that different automata may induce the same strategy; actually an adequate measure on the strategic complexity of an automaton has been shown to be its size, *i.e.* the number of its states $|Q^i|$ (see [Au81] and [Ney98]). Given a size $n \in \mathbb{N}^*$ we note $\mathcal{M}^i(n)$ the set of all automata of size n for player i; we have $|\mathcal{M}^i(n)| = n|A^i|^n n^{n|A^{-i}|}$.

3 Cyclicity of strategies and main result

3.1 Cyclicity of a strategy

Following Neyman ([Ney98]) a strategy σ^i of player i is said to be *compatible* with the play $\omega = ((a_1^1, a_1^2), \ldots, (a_t^1, a_t^2))$ (which will be denoted by $\sigma^i \rightsquigarrow \omega$) if

$$\forall l \in [\![1, t]\!],\ \sigma^i\left((a_1^1, a_1^2), \ldots, (a_{l-1}^1, a_{l-1}^2)\right) = a_l^i$$

Note that for any strategy σ^i of player i, any play ω, any action a^i and adverse actions $a^{-i} \neq b^{-i}$ we have $\sigma^i \rightsquigarrow (\omega, (a^i, a^{-i})) \Leftrightarrow \sigma^i \rightsquigarrow (\omega, (a^i, b^{-i}))$.

Moreover a play $\omega \in H_n$ is said to be *in correspondance for player i* (which will be denoted $\omega \in \boldsymbol{\mathcal{C}}_n^i$) if

$$\exists \phi : A^i \to A^{-i}\ / \forall t \in [\![1, n]\!],\ \omega_t^2 = \phi(\omega_t^1)$$

As will be discussed in the next paragraph implementing an efficient strategy against player one through an automaton requires an additional hypothesis be made on the strategies of player one, which relies on the following definition.

Definition 1 *A strategy σ^i is said to be T-*cyclic *for $T \geq 1$ if $\forall \omega \in H_\infty, \sigma^1 \rightsquigarrow \omega$,*

$$\left(\exists \alpha \in \boldsymbol{\mathcal{C}}_T^i, \alpha \lhd\!\!] \, \omega\right) \Rightarrow \left(\exists \beta \in H_n, \beta \lhd\!\!] \, \alpha, n \geq 1,\ / \, \sigma^1 \rightsquigarrow \omega\beta\beta\beta\ldots\right)$$

In such a case n is called the period *of σ^1 along the play ω (which depends on ω).*

Note that being T-cyclic implies being $(T+1)$-cyclic. For example the "tit-for-tat" strategy is 1-cyclic; however the "progressive tit-for-tat" is not.

The rational behind the definition of cyclicity lies in the following property :

Proposition 1 *The strategy induced by any automaton of size n is n-cyclic.*

which is a direct consequence of the definition of an automaton (see appendices for a proof).

The main result of this paper holds under assumptions that are formulated in terms of cyclicity; although they could have been expressed in terms of the size of some automata these weaker assumptions enlarge the application domain of the theorem to bounded recall strategies (as will be exposed in section 5).

3.2 Situation of the problem and main result

Let G be a finite two-players zero-sum game G and let us consider G_∞ the associated infinitely repeated game. For each profile of pure strategies $(\sigma^1, \sigma^2) \in S_\infty$ let

$$N\left(\sigma^1, \sigma^2\right) = \left|\left\{\tau \in \mathbb{N} \,/\, r\left(\omega\left(\sigma^1, \sigma^2\right)_\tau\right) > u^1(G)\right\}\right|$$

be the total number of stages at which player two does not beat player one, and

$$T\left(\sigma^1, \sigma^2\right) = \inf\left\{t_0 \in \mathbb{N} \,/\, \forall \tau > t_0 \,/\, r\left(\omega\left(\sigma^1, \sigma^2\right)_\tau\right) \leq u^1(G)\right\}$$

be the first stage after which player two always beats player one (with $\inf \emptyset = +\infty$).

Let $E \subset S^1_\infty$ be a non-empty set of allowed (pure) strategies for player one, and assume E is known by player two. Note however that player two does not know which strategy among E player one actually selects to play in G. Define for any $\sigma^2 \in S^2_\infty$

$$N\left(E, \sigma^2\right) = \sup_{\sigma^1 \in E} N(\sigma^1, \sigma^2) \quad \text{and} \quad T\left(E, \sigma^2\right) = \sup_{\sigma^1 \in E} T(\sigma^1, \sigma^2)$$

the total *number of failures* and the *time-to-win* of σ^2 respectively.

The purpose of this article is first to determine if there exists a strategy σ^2 for player two such that $N(E, \sigma^2)$ and $T(E, \sigma^2)$ are both finite (*i.e.* player two effectively beats player one), and more precisely to minimize the complexity of such a winning strategy as a function of the complexity of the strategies of E. The second goal is to minimize $N(E, \sigma^2)$ and $T(E, \sigma^2)$. In this article complexity is tackled with in terms of size of finite automata and in terms of size of the memory of a bounded recall strategy.

Under this context the following results have already been proven :

Theorem 1 (Ben-Porath 1993) *Let $n \in \mathbb{N}^*$. Then there exists $M^2 \in \mathcal{M}^2\left(n|\mathcal{M}^1(n)|\right)$ such that $\forall M^1 \in \mathcal{M}^1(n)$ $r_\infty(\sigma(M^1), \sigma(M^2)) \leq u^1(G)$*

In other words setting $E = \mathcal{M}^1(n)$ there exists $M^2 \in \mathcal{M}^2\left(n|\mathcal{M}^1(n)|\right)$ such that $N(E, \sigma(M^2)) < +\infty$ (and thus $T(E, \sigma(M^2)) < +\infty$). Note that $n|\mathcal{M}^1(n)| = n^{Cn+1}$ for $C = |S_2| + \log|S_1| + 1$ as was claim in the introduction of this paper.

Moreover although it is not explicitely stated it follows from the proof of Theorem 1 that $T(E, \sigma(M^2)) = N(E, \sigma(M^2)) = n(|E| - 1)$. [6]

The core of the construction of M^2 consists for player two in assuming that it faces a certain automaton M^1; as soon as player one plays an action that M^1 would not have played, player two *chooses* (through an arbitrary ordering among automata) another compatible automaton and assumes it plays against it, and so on until its actual opponent is fully identified (see [Por93] for details). The main difference of the strategy suggested in

[6] Following [Por93] notations let $N = |\mathcal{M}^1(n)|$ and let $\alpha(1), \ldots; \alpha(N)$ be the (finite) enumeration of all possible automata for player one, and consider $M^1 = \alpha(N)$. Then M^2 will need to successively switch from the best-reply automata $\overline{\alpha(1)}, \ldots, \overline{\alpha(N-1)}$ and possibly each of their n copies before eventually switching to $\overline{\alpha(N)}$, and hence the result.

[NO00] (apart from the fact that it is not directly implemented by an automaton) is that it updates its hypothesis on its opponent in a maximum likelihood fashion, which leads to a much stronger victory of player two as stated in the following result :

Theorem 2 (Neyman-Okada 2000) *There exists* $\sigma^2 \in S^2_\infty$ *such that* $N(E, \sigma^2) \leq \lceil \log_2 |E| \rceil$.

The proof of this theorem is based on the following observation : passed histories provide for a partition of the set E of strategies of player one. More precisely for any stage $t \in \mathbb{N}$ and any history $h_t \in H_t$ define the partition $\mathcal{I}_t = \bigcup_{h_t \in H_t} C_t(h_t)$. $\mathcal{I}_t$ materializes the information set of player two at stage t, since after h_t player two knows that the strategy of player one is in $C_t(h_t)$. The desired property is that $\mathcal{I}_t$ converges to a *fully identified partition*, *i.e.* an union of singletons (or more adequately to a union of subsets such that all strategies of a given subset are equivalent). The point here is to set player two's action after history h_t to be the best-reply to the action played by the largest subset $\mathcal{I}_{t+1}(h_t, \cdot)$ *at stage* $t+1$. [7] This ensures that at next stage, at most half of the remaining compatible strategies of E will avoid being beaten, and hence the result.

However σ^2 has an unspecified complexity in terms of finite automata or bounded-recall; moreover $T(E, \sigma^2)$ is unbounded. The purpose of the next section is to implement σ^2 through an automaton and to bound $T(E, \sigma^2)$. Note that this requires that the convergence of $\mathcal{I}_t$ to a fully-identified partition happens in uniformly bounded time, which is not true in general without further assumptions on E. The next sections shows that assuming an uniformly bounded cyclicity of the strategies of E is sufficient, as stated in the following result :

Theorem A *Let* $T \in \mathbb{N}^*$ *and suppose that* $E \subset S^1_\infty$ *is such that each* $\sigma^1 \in E$ *is* T*-cyclic. Then there exists* $M^2 \in \mathcal{M}^2(2T|E|)$ *such that* $T(E, \sigma(M^2)) \leq 2T \lceil \log_2 |E| \rceil$ *and* $N(E, \sigma(M^2)) \leq \lceil \log_2 |E| \rceil$.

This result is proven in the next section. The rational of stating the hypothesis on E in terms of cyclicity is that it allows us to repharse Theorem A in terms both of finite automata and of bounded-recall strategies, as will be done in section 5.

These results may be summarized as follows:

	E. Ben-Porath	A. Neyman & D. Okada	This paper
Strategies of player one	Automaton $E = \mathcal{M}^1(n)$	*any* (unbounded complexity)	strategies which are T-cyclic
Strategy of player two	Automaton of size $n\|E\|$	functional form (unspecified complexity)	Automaton of size $2T\|E\|$
Total number of failures	$(\|E\| - 1)$	$\lceil \log_2 \|E\| \rceil$	$\lceil \log_2 \|E\| \rceil$
Time-to-win	$n(\|E\| - 1)$	*unbounded*	$2T \lceil \log_2 \|E\| \rceil$

[7] If there are more than one subset $\mathcal{I}_{t+1}(h_t, \cdot)$ of maximum cardinality, anyone will do.

4 Proof of Theorem A

Let $T \in \mathbb{N}^*$ and $E \subset S^1_\infty$ be such that each $\sigma^1 \in E$ is T-cyclic. We built an automaton $M^2 \in \mathcal{M}^2(2T|E|)$ such that $T(E, \sigma(M^2)) = 2T\lceil \log_2 |E| \rceil$. The sketch of the proof is as follows. First we build the automaton M^2 which implements the desired winning strategy; then we prove that it is of adequate size and effectively beats any $\sigma^1 \in E$ sufficiently quickly.

The construction of M^2 relies on two distinct ideas : on the one hand the information set $\mathcal{I}_t$ of [NO00] is explicitely materialized through an automaton; on the other hand cyclicity ensures that a finite number of states may be sufficient to implement such an automaton. In the second part of the proof both ideas ensure that $T(E, \sigma(M^2)) \le 2T\lceil \log_2 |E| \rceil$ (each accounting for the $\lceil \log_2 |E| \rceil$ or for the $2T$ factors respectively); eventually cyclicity ensures that the size of M^2 is bounded by $2T|E|$.

4.1 Construction of M^2

Let $b_0 \in A^2$ be a fixed action for player two. For any stage $t \in \mathbb{N}^*$ and any $\omega \in H_t$ let $C_t(\omega) = \{\sigma \in E \ / \ \sigma \rightsquigarrow \omega\}$ be the set of all compatible strategies. We build an infinite tree $(Q_k, f^2_k)_{k\in\mathbb{N}}$ whose nodes $(Q_k)_{k\in\mathbb{N}}$ materialize the informations sets $(\mathcal{I}_k)_{k\in\mathbb{N}}$ and whose edges are defined by $(f^2_k)_{k\in\mathbb{N}}$. Together with adequate actions $(g^2_k : Q_k \to A^2)_{k\in\mathbb{N}}$ this will be the ground of M^2. To simplify the presentation we also introduce the set of all encountered histories $(\Omega_k)_{k\in\mathbb{N}}$. Let us define $\mathcal{T} = (\Omega_k, Q_k, f^2_k, g^2_k)_{k\in\mathbb{N}}$ recursively in the following way :

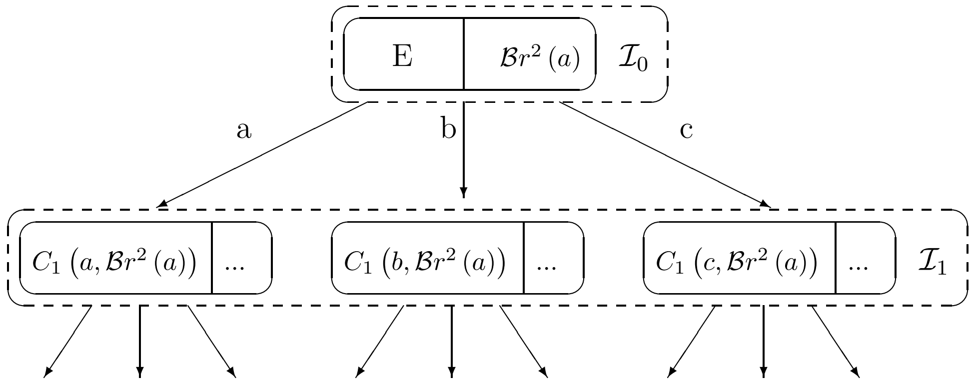
- Let $\Omega_0 = \{\oslash\}$ and $Q_0 = C_0(\Omega_0) = \{E\}$.
- Let $k \ge 0$ be such that Ω_k and Q_k are well-defined; let us define g^2_k, f^2_k and Ω_{k+1}.

  Let $g^2_k : \begin{pmatrix} Q_k & \to & A^2 \\ C_k(\omega) & \mapsto & \mathcal{B}r^2\left(argmax_{a^1\in A^1}|C_{k+1}(\omega,(a,b_0))|\right) \end{pmatrix}$ be the best reply to the "most played" action of player one.

  Let $\Omega_{k+1} = \{(\omega,(a,b)) \ / \ \omega \in \Omega_k \,\text{and}\, a \in A^1 \,\text{and}\, b = g^2_k(C_k(\omega))\}$ be the set of all possible histories *given the action of player two at the last stage*, and let $Q_{k+1} = C_{k+1}(\Omega_{k+1})$.

  Then define $f^2_k : \begin{pmatrix} Q_k \times A^1 & \to & Q_{k+1} \\ (C_k(\omega), a^1) & \mapsto & C_{k+1}(\omega,(a,b_0)) \end{pmatrix}$ the transition function that goes from the state $C_k(\omega)$ to the state $C_{k+1}(\omega,(a,b_0))$. Note that $C_{k+1}(\omega,(a,b_0)) = C_{k+1}(\omega,(a,g^2_k(C_k(\omega))))$ so that $C_{k+1}(\omega,(a,g^2_k(C_k(\omega))))$ is effectively the next state.

The infinite tree $\mathcal{T}$ thus has an infinite number of rows whose nodes $C_k(\omega)$ form for each $\omega \in \Omega_k$ a partition of E, which materializes $\mathcal{I}_k$, as illustrated in the following :



In addition at each node $C_k(\omega)$ the action $g_k^2(C_k(\omega))$ is defined to be the best-reply to the "most-often played" action (namely a in the above illustration). In this example if player one plays a, then he will be beaten at stage one and player two knows that $\sigma_1 \in C_1(a, \mathcal{B}r^2(a))$. Otherwise player one plays for example b and although may get a higher payoff than $u^1(G)$ at stage one, on the other hand player two knows that $\sigma_1 \in C_1(b, \mathcal{B}r^2(a))$. The key point here is that $|C_1(b, \mathcal{B}r^2(a))| \leq \frac{1}{2}|E|$ since $|C_1(b, \mathcal{B}r^2(a))| \leq |C_1(a, \mathcal{B}r^2(a))|$.

There are two caveats that need however to be addressed : first $\mathcal{T}$ has an *infinite* number of nodes; and second it may happen that no two different actions are possibly played from one stage to another (so that $(\mathcal{I}_t)$ need not necessarily converge to a fully-identified partition). In order to tackle with these issues we need the following definitions.

For $k \geq 2T$ let

$$W_k = \{\omega = ((a_1, b_1), \ldots, (a_k, b_k)) \in \Omega_k \;/\; \forall l \in [\![k - 2T + 1, k]\!],\; b_l = \mathcal{B}r^2(a_l)\}$$

be the set of the plays have been victorious during the last $2T$ stages. Define $\overline{W}_k = W_k \setminus \bigcup_{l=2T}^{k-1} W_l \times H_{k-l}$. The point here is that due to the following result it is not necessary to wait for all strategies among a given branch of $\mathcal{T}$ to differentiate one from another (which may actually not always happen) :

Lemma 1 *Let $k \geq 2T$ and $\omega \in H_k$ be such that $(\omega_{k-2T+1}, \ldots, \omega_k)$ is a best-reply path. Then $\forall n \in \mathbb{N}$, $\forall \alpha^1 \in (A^1)^n$, $\forall \sigma^1, \mu^1 \in C_k(\omega)$, $\sigma^1 \rightsquigarrow (\omega\underline{\alpha^1}) \Leftrightarrow \mu^1 \rightsquigarrow (\omega\underline{\alpha^1})$*

This lemma (which is proven in appendices) ensures that once a group of strategies have been beaten $2T$ times they will *never* distinguish one from each other (as long as player two plays a best reply at each stage). It is thus sufficient to beat one of them; but since they are T-cyclic, it is sufficient to *loop-back* n stages earlier where $n \leq T$ is the period of one of them along ω (as will be shown below).

Let us define the automaton M^2 explicitly.

Let $N = 2T\lceil \log_2 |E| \rceil$ and let $Q = Q_1 \bigcup \cdots \bigcup Q_{2T-1} \bigcup C_{2T}(\overline{W}_{2T}) \bigcup \cdots \bigcup C_N(\overline{W}_N)$. Q is the smallest subset of $(Q_k)_{k \in [\![1,N]\!]}$ that contains all nodes corresponding to a winning history in W_k. Let $g : \begin{pmatrix} Q & \rightarrow & A^2 \\ C_k(\omega) & \mapsto & g_k^2(C_k(\omega)) \end{pmatrix}$. Now define the transition function f

step-by-step : first implement all transitions already defined on $(Q_k)_k$ through f_k^2, and then add all loop-back transitions from the leaves $\overline{W}_k$ after which player one will get beaten at each stage. Formally for any ω in a given E let $\tau(\omega)$ be be an arbitrary selection, and let $n(\omega) \leq T$ be its period along ω.[8] Define

$$f : \left(\begin{array}{ccc} Q \times A^1 & \rightarrow & Q \\ (C_k(\omega), a) & \mapsto & \left\{ \begin{array}{ll} f_k^2(C_k(\omega), a) & \text{if } \omega \notin \overline{W}_k \\ C_{k-n(\omega)}(\omega_{|k-n(\omega)}) & \text{otherwise} \end{array} \right. \end{array} \right)$$

This completes the definition of the automaton $M^2 = [Q, q_* = E, f, g]$.

4.2 Proof of the adequation of M^2

We now show that M^2 is a quickly winning automaton. Let $\sigma^1 \in E$ be a strategy for player one, and set $\omega = \omega\left(\sigma^1, \sigma\left(M^2\right)\right)$. First let us show the following lemma :

Lemma 2 *Suppose that there exists $k \in \mathbb{N}^*$ such that $\omega_{|k} \in \overline{W}_k$. Then σ^1 will be beaten by $\sigma(M^2)$ at all stages after the first k ones.*

Proof Let $k \in \mathbb{N}^*$ be such that $\omega_{|k} \in \overline{W}_k$. At stage $k+1$ player two loops back to node $C_{k-n(\omega)+1}(\omega_{|k-n(\omega)+1})$ and plays a best reply to $\tau(\omega)$ at stage $k - n(\omega) + 1$; note that by definition of $n(\omega)$ this is also a best reply to $\tau(\omega)$ at stage $k+1$. In addition we have $\tau(\omega) \rightsquigarrow \omega_{|k}\alpha\alpha\alpha\ldots$ for $\alpha = (\omega_{k-n(\omega)+1}, \ldots, \omega_k)$. But since $\omega \in \overline{W}_k$ and $n(\omega) \leq T < 2T$, α is a best-reply path and lemma 1 applies; thus $\sigma^1 \rightsquigarrow \omega_{|k}\alpha\alpha\alpha\ldots$ since $\sigma^1 \in \overline{W}_k$, and in particular σ^1 plays the same action as $\tau(\omega)$ at stage $k+1$. Thus player two plays a best-reply to it at that stage. But then in turn the same reasoning proves that both $\tau(\omega)$ and σ^1 play the same action at stage $k+2$, and thus that player two plays a best-reply at stage $k+2$. Then a simple induction shows that $\sigma(M^2) \rightsquigarrow \omega_{|k}\alpha\alpha\alpha\ldots$, so that $\omega = \omega_{|k}\alpha\alpha\alpha\ldots$: player 1 will be beaten by player two at all stages after the first k ones.

□

Now define $p_0 = 0$ and for $l \geq 1$ $p_l = \inf\{p > p_{l-1} \;/\; r\left(\omega_p\right) > u^1(G)\}$ be the l-th stage at wich player one does not lose (with $\inf \emptyset = +\infty$). We have

Proposition 2 $\forall l > \log_2 |E|,\ p_l = +\infty$.

Proof Assuming $p_{l+1} < +\infty$ we have $|C_{p_{l+1}}(\omega_{|p_{l+1}})| \leq \frac{1}{2}|C_{p_l}(\omega_{|p_l})|$; thus for all $l \in \mathbb{N}$

$$p_l < +\infty \quad \Rightarrow \quad 1 \leq |C_{p_l}(\omega_{|p_l})| \leq \frac{1}{2^l}|E|$$

Since $\frac{1}{2^l}|E| < 1$ for $l > \log_2 |E|$ the result follows.

□

[8]The point here is that $n(\omega)$ is the period of one (undetermined) strategy $\tau(\omega)$ of $C_k(\omega)$, although the transition is consistant for any $\tau \in \mathbb{C}_k(\omega)$ due to lemma 1. This is shown in the next paragraph.

Note that this also proves that $N\left(E,\sigma\left(M^2\right)\right) \leq \lceil \log_2 |E| \rceil$: player one can not avoid being beaten more that $\lceil \log_2 |E| \rceil$ times. And besides there always exists a $k \in \mathbb{N}^*$ such as in lemma 2, and and it is bounded by $2T\lceil log_2|E|\rceil$:

Proposition 3 *There exists $k \leq 2T\lceil log_2|E|\rceil$ such that $\omega_{|k} \in \overline{W}_k$.*

Proof On the one hand because of lemma 2 we have $\forall l < l'$, $p_{l'} < p_l + 2T$. On the other hand because of proposition 2 we have $\{p > p_{\lceil \log_2 |E| \rceil} \in \mathbb{N} \ / \ r(\omega_p) > u^1(G)\} = \emptyset$. But since $p_l \geq p_0 + l \cdot (2T)$ this means that $\{p > 2T\lceil log_2|E|\rceil \in \mathbb{N} \ / \ r(\omega_p) > u^1(G)\} = \emptyset$, and hence the result.

□

It follows from proposition 3 and lemma 2 that M^2 beats any strategy σ^1 of E at all stages after the first $2T\lceil log_2|E|\rceil$ ones. In other words we have $T\left(E,\sigma\left(M^2\right)\right) \leq 2T\lceil log_2|E|\rceil$.

To complete the proof of Theorem A observe that the size of M^2 is no larger than $2T|E|$: on the one hand there are at most $|E|$ leaves (nodes from which the automaton loops back) in the transition graph of the automaton since each leaf contains at least one strategy. And on the other hand, any node belongs to at least one winning path of length at most $2T$ (since at each node the automaton plays the best reply to at least one strategy, and beating a strategy requires at most $2T$ steps before looping back). Thus the transition graph of the automaton contains at most $2T|E|$ nodes, that is $|Q| \leq 2T|E|$.

5 Application to finite automata and bounded-recall strategies

5.1 Application to finite automata

In the context of finite automata since the strategy induced by an automaton is cyclic Theorem A leads to the following result :

Corollary 1 *There exists an automaton of size $2n|\mathcal{M}^1(n)|$ which beats any automaton of size at most n at each stage except possibly at the first $2n\lceil \log_2 |\mathcal{M}^1(n)|\rceil$.*

This is a direct consequence of Proposition 1. This result is similar to Theorem 1. Let N^2 be the automaton defined in [Por93] and let us compare N^2 with M^2. On the one hand M^2 may be twice larger than N^2; recall however that the size of N^2 is n^{Cn+1}. And on the other hand we have $N\left(E,\sigma\left(N^2\right)\right) = (|E|-1)$ whereas $N\left(E,\sigma\left(M^2\right)\right) = 2n\lceil \log_2 |E| \rceil$ which is logarithmically smaller, and in the same way $T\left(E,\sigma\left(N^2\right)\right) = n\left(|E|-1\right)$ whereas $T\left(E,\sigma\left(M^2\right)\right) = \lceil \log_2 |E| \rceil$. This is a consequence of the fact that instead of updating his beliefs on the strategy of player one arbitrarily as does [Por93], in Theorem A player two does it in a maximum likelihood fashion as suggested in [NO00].

Note further that Theorem 2 addresses the case where the strategies of player one are completely arbitrary - as long as E is bounded. However this only provides a bound on $N(E,\sigma^2)$ since $T(E,\sigma^2)$ can not be bounded in terms of $|E|$ only : consider for any $N \in \mathbb{N}$ the case where E contains only two strategies which unconditionnally play the same until stage N but then differentiate one from another. As a consequence without further assumptions on E there can be no automaton whose size would depend only on $|E|$ which would eventually beat all strategies of E : it is required that any two strategies have bounded differenciation time, *i.e.* once they have played the same for a sufficiently long time they are indeed equivalent. In this sense Theorem A establishes a stronger result based on a stronger hypothesis; the point here is that this hypothesis is stated in terms of cyclicity. The main advantage is that it is related to each strategy *individually* instead of E globally. [9] It is the essence of Lemma 1 to turn the individual cyclicity into the global property of bounded differenciation time.

5.2 Extension to bounded-recall strategies

Apart from finite state automata another model of bounded rationality has been defined by E. Lehrer in the mean of *Stationary Bounded Recall Strategies* (SBRS) :

Definition 2 *A strategy σ^i for player i is said to be n-SBRS if there exists $e \in H_n$ and $\phi^i : (H_n \to A^i)$ such that $\forall t \in \mathbb{N}, \forall h \in H_t$, $\sigma^i_t(h) = \phi^i\left(h_{t-n+1}, \ldots, h_t\right)$ with the convention that $(h_{-n}, \ldots, h_{-1}) = (e_1, \ldots, e_n)$. In this case we note $\sigma^i = \sigma\left(\phi^i\right)$.*

This states that the action played by σ^i at stage t depends only on the last n actions, provided an initial memory e. Note that ϕ^i might be an arbitrarily complex function as long as it remains deterministic. Setting $l = |A|$ the following result is proven in [Leh94] :

Theorem 3 *There exists a $\left(l^2+1\right)^n$-SBRS strategy of player two that beats any n-SBRS strategy of player one after stage $\left(l^2+1\right)^n + l^n$.*

The aim of this paragraph is to elaborate a similar result on the basis of Theorem A . Note that a direct consequence of the definition of a SBRS strategy is :

Proposition 4 *Any n-SBRS strategy is l^n-cyclic.*

As a consequence setting $k = |A^2|$ we have

Corollary 2 *There exists an automaton of size $2l^n k^{l^n}$ for player two that beats any n-SBRS strategy at any stage except possibly at the first $2l^{2n}\lceil \log_2 k\rceil$ ones.*

Proof Let E be the set of all n-SBRS strategies; we have $|E| = k^{l^n}$. Because of Proposition 4 any strategy of E is l^n-cyclic. Thus by Theorem A let M^2 be an automaton of size $2T|E| = 2l^n k^{l^n}$ such that $N\left(E, \sigma\left(M^2\right)\right) = \lceil \log_2 |E| \rceil$ and $T\left(E, \sigma\left(M^2\right)\right) = 2T\lceil \log_2 |E| \rceil$. We have $T\left(E, \sigma\left(M^2\right)\right) = 2l^n \left\lceil \log_2 \left(k^{l^n}\right)\right\rceil \leq 2l^{2n}\lceil \log_2 k\rceil$, and hence the result.

□

[9] With the provision that there is an *uniform* bound on the latest stage after which a strategy is cyclic.

Note that similarly to what happens in the case of finite automata, the number of failures is roughly logarithmically smaller than the minimum complexity of the strategy of player two (be it the size of an autoaton or of a memory) : this is an intrisic property of Theorem A . However the interest of Corollary 2 is moderated by the huge size of M^2 : even in the simple case where $|A^1| = |A^2| = 2$ it states that any 5-SBRS can be beaten by a given automaton of size $\ldots 2^{1035}$!

This is why we rather consider the case where both player one and player two play bounded-recall strategies, as in the following result :

Theorem B *Let $T \in \mathbb{N}^*$ and suppose that $E \subset S^1_\infty$ is such that each $\sigma^1 \in E$ is T-cyclic. Then there exists a $2T\lceil \log_2 |E| \rceil$-SBRS for player two which beats any cyclic strategy of E at each stage except possibly at the first $2T\lceil \log_2 |E| \rceil$ ones.*

Proof The proof of Theorem A can be tailored to fit the case where both players play bounded-recall strategies. Obviously E is non-empty; let $T = l^n$. Let us define ϕ^2 such that $T(E, \sigma(\phi^2)) \leq 2T\lceil \log_2 |E| \rceil$. Recall that there are indeed two operations involved in Theorem A : first to progressively identify the strategy σ^1 actually selected by player one, and then to play a best-reply action at each stage.

In order to implement the progressive identification of player one's strategy let $\mathcal{T} = (\Omega_k, Q_k, f^2_k, g^2_k)_{k \in \mathbb{N}}$ be the infinite information tree defined in the proof of Theorem A , and let M^2 be the corresponding automaton. The point here is to define ϕ^2 based on a finite subtree of $\mathcal{T}$ so that it plays like M^2; this relies on the following definition.

Definition 3 *A play $\omega \in H_n$, $n \in \mathbb{N}^*$ is said to be t-*looping *at stage $\tau < n$ if*
$\exists \alpha \in H_{\tau-1}, \exists \lambda \in H_p, \exists \beta \in H_r$, $p \geq 1$ / $\omega = \alpha\lambda\lambda\beta$ *and* $\beta \llbracket\!< \lambda$ *and* $2p + r \geq t$.

Let $N = 2T\lceil \log_2 |E| \rceil$. The proof consists in defining a N-SBRS ϕ^2 such that the play of any σ^1 against $\sigma(\phi^2)$ is looping at stage $2T\lceil \log_2 |E| \rceil$ at the latest.

Let $\phi^2 : \left(\begin{array}{ccc} \bigcup_{n \leq N} H_n & \rightarrow & A^2 \\ \omega & \mapsto & \left\{ \begin{array}{ll} g^2_n(C_n(\omega)) & \text{if } \omega \in H_n \text{ is not } 2T\text{-looping} \\ \mathcal{B}r^2\left(\omega^1_{r+1}\right) & \text{if } \omega \text{ is } 2T\text{-looping} \end{array} \right. \end{array} \right)$. Note that in the case where ω is looping the pair of actions $\omega_{r+1} \in \{\omega_1, \ldots, \omega_r\}$ has already been played so that the definition of ϕ^2 always makes sense.

Lemma 3 *$\sigma(\phi^2)$ and $\sigma(M^2)$ are equivalent against all strategies of E :*
$\forall \sigma^1 \in E,\ \omega(\sigma^1, \sigma(\phi^2)) = \omega(\sigma^1, \sigma(M^2))$.

Proof Let $\sigma^1 \in E$, and set $\omega = \omega(\sigma^1, \sigma(\phi^2))$ and $w = \omega(\sigma^1, \sigma(M^2))$. Because of Proposition 3 let $k \leq N$ be the smallest stage such that $w_{|k} \in \overline{W}_k$ ($k \geq 2T$ by definition of $\overline{W}_k$). It follows that w is $2T$-looping. Let us prove by induction that $\forall t \in \mathbb{N}$, $w_{|t} = \omega_{|t}$.

- At $t = 0$ we have $w_{|0} = \oslash = \omega_{|0}$.

- Suppose that there exists $t \geq 1$ such that $w_{|t-1} = \omega_{|t-1}$, and let us prove that $w_t = \omega_t$. Since $w_t^1 = \sigma^1(w_{|t-1}) = \sigma^1(\omega_{|t-1}) = \omega_t^1$ it remains to be proven that $w_t^2 = \omega_t^2$.

  Suppose first that $t-1 < k$; then $w_{|t-1} \notin W_{t-1}$ and thus $\omega_{|t-1} = w_{|t-1}$ is not $2T$-looping. Thus
  $$w_t^2 = g_{t-1}^2\left(C_{t-1}\left(w_{|t-1}\right)\right) = g_{t-1}^2\left(C_{t-1}\left(\omega_{|t-1}\right)\right) = \omega_t^2$$
  (recall that $k \leq N$ so that $\omega_{|t-1}$ fits into the bounded recall H_N of ϕ^2).

  Otherwise, $t-1 \geq k$ and thus $w_{|t-1} = \omega_{|t-1}$ is looping, and let $\alpha \in H_n, \lambda \in H_p, \beta \in H_r$ be such that $\omega_{|t-1} = \alpha\lambda\lambda\beta$ with $\beta \Subset \lambda$. The caveat here is that the whole $\omega_{|t-1}$ may not fit into the bounded recall H_N of ϕ^2. If $t-1 \leq N$ then this is not the case and
  $$\omega_t^2 = \mathcal{B}r^2\left(\omega_t^1\right) = \mathcal{B}r^2\left(\omega_{t-p}^1\right) = \mathcal{B}r^2\left(w_{t-p}^1\right) = \mathcal{B}r^2\left(w_t^1\right) = w_t^2$$
  Otherwise $t-1 > N$ and $\omega_t^2 = \phi^2\left(\omega_{t-N}, \ldots, \omega_{t-1}\right)$. But since ω_t is looping, $\omega_t^2 = \phi^2\left(\omega_{t-N}, \ldots, \omega_{t-1}\right)$ is a best-reply path; and since σ^1 is T-cyclic we have $p \leq T$. As a consequence since $N = 2T\lceil\log_2|E|\rceil \geq 2T$, $\omega_{t-N}, \ldots, \omega_{t-1}$ is also looping. Thus similarly
  $$\omega_t^2 = \mathcal{B}r^2\left(\omega_t^1\right) = \mathcal{B}r^2\left(\omega_{t-p}^1\right) = \mathcal{B}r^2\left(w_{t-p}^1\right) = \mathcal{B}r^2\left(w_t^1\right) = w_t^2$$
  which completes the induction.

This shows that $\sigma(\phi^2)$ and $\sigma(M^2)$ are equivalent against σ^1.

□

To complete the proof of Theorem B observe that ϕ^2 has a bounded-recall of size $N = 2T\lceil\log_2|E|\rceil$ as claimed; moreover we have $T\left(E, \sigma\left(\phi^2\right)\right) = T\left(E, \sigma\left(M^2\right)\right) \leq 2T\lceil\log_2|E|\rceil$ by Theorem A . Note that in addition $N\left(E, \sigma\left(\phi^2\right)\right) = N\left(E, \sigma\left(M^2\right)\right) \leq \lceil\log_2|E|\rceil$.

□

A direct consequence of Theorem B is the following :

Corollary 3 *There exists a $l^{2n}\left(\lceil\log_2 k\rceil + 1\right)$-SBRS strategy for player two which beats any n-SBRS strategy of player one at each stage after stage $l^{2n}\left(\lceil\log_2 k\rceil + 1\right)$.*

To complete the panorama we shall also mention the following :

Corollary 4 *There exists a $2n\lceil\log_2|E|\rceil$-SBRS strategy which beats any automaton of size n of player one at each stage after stage $2n\lceil\log_2|E|\rceil$.*

These result may be summarized in the following way :

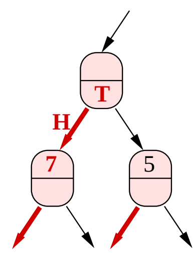
	player two implements his strategy through an automaton	player two plays a bounded-recall strategy
To beat any automaton of size n of player one	of size $2n\|E\|$ and wins after stage $2n\lceil\log_2\|E\|\rceil$	of size $2n\lceil\log_2\|E\|\rceil$ and wins after stage $2n\lceil\log_2\|E\|\rceil$
To beat any n-SBRS strategy of player one	of size $2l^n k^{l^n}$ and wins after stage $2l^{2n}\lceil\log_2 k\rceil$	of size $l^{2n}\left(\lceil\log_2 k\rceil + 1\right)$ and wins after stage $l^{2n}\left(\lceil\log_2 k\rceil + 1\right)$

6 Example

To illustrate the main result we consider in this section the infinitely repeated "*Matching Pennies*" zero-sum game defined by $A^i = \{H^i, T^i\},\ i \in \{1, 2\}$ and with payoff r given by

	H^2	T^2
H^1	+1	-1
T^1	-1	+1

Let us consider the case where player one is restricted to strategies that can be implemented by an automaton of size two : $E = \mathcal{M}^1(2)$. To build the complete automaton M^2 which eventually beats any strategy of E let us label each edge by the action played by player one and label each state $C(\omega)$ both by its cardinality $|C(\omega))|$ and by the associated action $g^2(C(\omega))$, as in the following example :



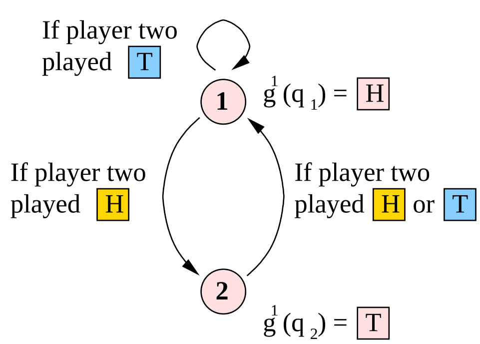
By definition the action played by player two at the source node is always the best reply to the edge which leads to the node with highest cardinality (by convention all those edges will be put left-most). In this situation at the next stage 7 automata would play H and 5 would play T; player two thus plays $T = \mathcal{B}r^2(H)$.

In addition fully identified automata are represented below their corresponding identification leaf. Observe that automata for player one may not have more than two states and may not play more than two actions; thus we suppose that $g^1 : \begin{cases} q_1 \mapsto & \text{H} \\ q_2 \mapsto & \text{T} \end{cases}$. An automaton for player one is thus characterized by its transition function $f^1 : Q^1 \times A^2 \to Q^1$.

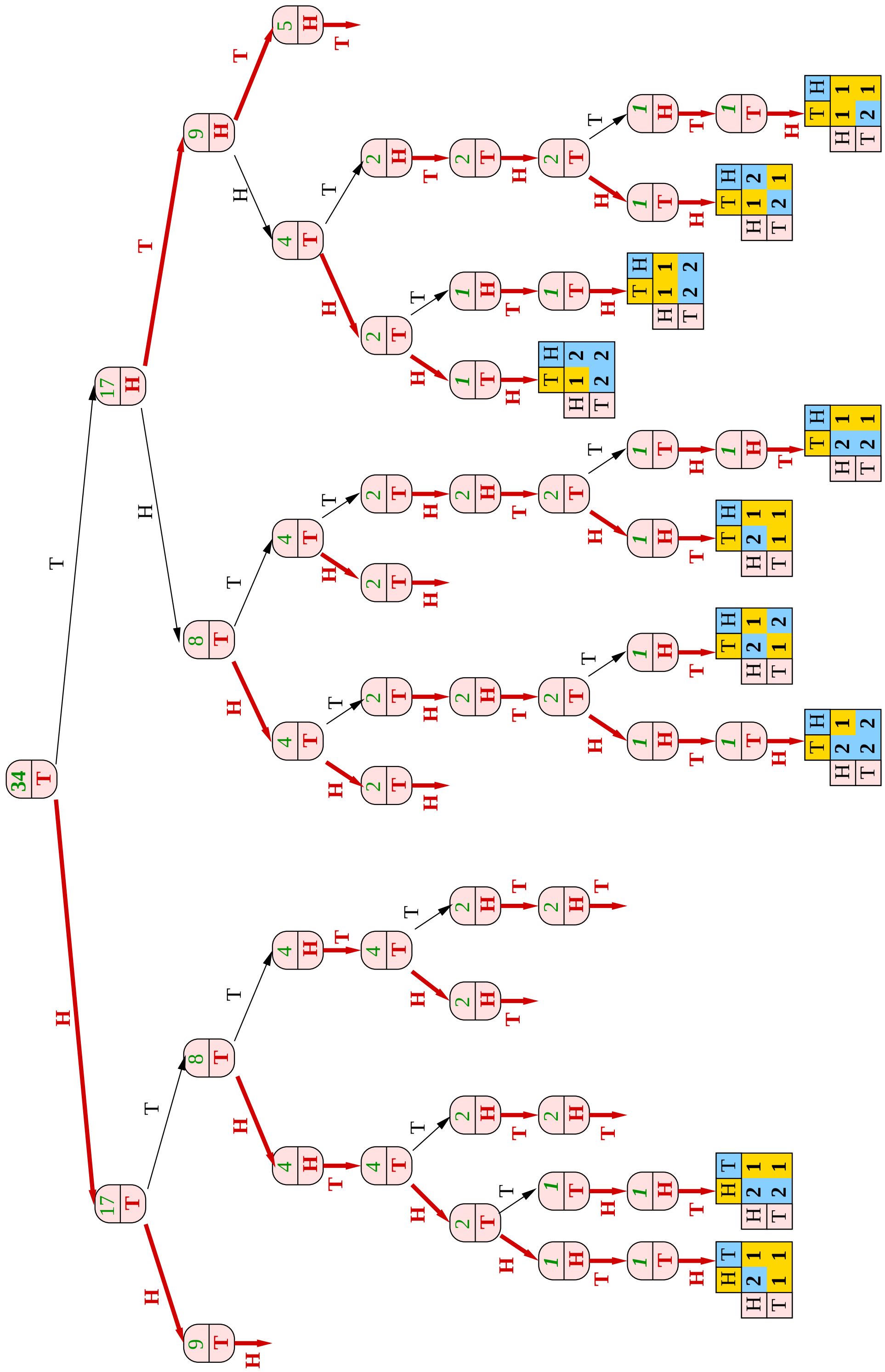
For example

	H	T
H	**2**	**1**
T	**1**	**1**

stands for the automaton



The following automaton for player two that beats all 34 non-equivalent automata of player one at any stage except possibly at the first 7 ones. This automaton is of size $49 \leq 2 \times 2 \times \mathcal{M}^1(2) = 4 \times (2 \cdot 2^2 \cdot 2^{2 \cdot 2}) = 512$.



7 Conclusion and future work

The results presented in this paper suggest that the learning tree structure presented in [NO00], provided an adequate hypothesis on the strategies of player one, constitue a convenient framework to answer questions (i), (ii) and (iii) whenever rationality bounds are expressed in terms of finite state automata or of bounded-recall strategies.

Practical analysis shows however that the theoritical bounds may be oversized : this is largely due to the fact that since player two plays a best-reply most of the time during the learning phase most of the strategies of player one play equivalently although they are not equivalent in general. This leads to a much quicker identification and a reduction of the number of nodes of $\mathcal{T}$ that are actually required; it remains unclear at this stage of this reduction can be quantified.

Besides in this paper only pure strategies are considered. On the first hand the results still hold in the case where player one plays a mixt strategy of the repeated game, which consists in randomly choosing at stage zero a pure strategy for the whole game. On the other hand if player one plays a behavioural strategy the identification tree $\mathcal{T}$ can not identify his strategies anymore; an approach based on coordination as studied in [GH03] seems more appropriate in this case. And in the case where player two in turn randomizes the discussion by E. Kalai and E. Solan ([KS00]) suggests that he may beat player one more efficiently would he implement an automaton with random transitions, although this remains unclear under the present context.

Acknowledgements

The authors wish to thank S. Sorin and J. Abdou for their help in the works presented in this paper. We also would like to thank O. Gossner and T. Tomala for their valuable remarks and suggestions during the preparation of this paper.

References

[Au81] R. Aumann, 1981. *Essays in Game Theory and Mathematical Economica in Honor of Oskar Morgenstern* , Bibliographisches Institut, Zürich pp $11-42$.

[Au97] R. Aumann, 1997. *Rationality abd Bounded Rationality* , Games and Economic Behavior, vol. 21 pp $2-14$.

[Por93] E. Ben Porath, 1993. *Repeated Games with Finite Automata* , Journal of Economic Theory, vol. 59 pp $17-32$.

[GH03] O. Gossner and Penélope Hernàndez, 2003. *On the compexity of coordination* , Mathematics of Operations Research, vol. 28 pp $127-141$.

[HU79] J. E. Hopcroft and J. D. Ullman, 1979. *Introduction to automata theory, languages, and computation* , Addison Wesley.

[Kal93] E. Kalai, 1993. *Bounded rationality and Strategic Complexity in Repeated Games* , Game Theory and Applications pp 131 − 157 .

[KS00] E. Kalai and E. Solan, 2000. *Randomization and Simplification* , Northwestern University, Center for Mathematical Studies in Economics and Management ScienceDiscussion paper 1283.

[Leh94] E. Lehrer, 1994. *Finitely many Players with Bounded Recall in Infinitely Repeated Games* , Games and Economic Behavior, vol. 7 pp 390 − 405 .

[Ney85] A. Neyman, 1985. *Bounded Complexity Justifies Cooperation in the Prisoneer's Dilemma* , Economic Letters, vol. 19 pp 227 − 229 .

[Ney98] A. Neyman, 1998. *Finitely Repeated Games with Finite Automata* , Mathematics of Operations Research, vol. 23 pp 513 − 551 .

[Ney97] A. Neyman, 1997. *Cooperation, Repetition and Automata* , Cooperation : Game Theoretic Approaches, *Sergiu Hart and Andreu Mas-Colell, Editors*, Springer-Verlag 1997.

[NO00] A. Neyman and D. Okada, 2000. *Two Person Repeated Games with Finite Automata* , International Journal of Games Theory, vol. 19.

[Rub88] A. Rubinstein, 1988. *The Structure of Nash Equilibria in Repeated Games with Finite Automata* , Econometrica, vol. 56 pp 1259 − 1281 .

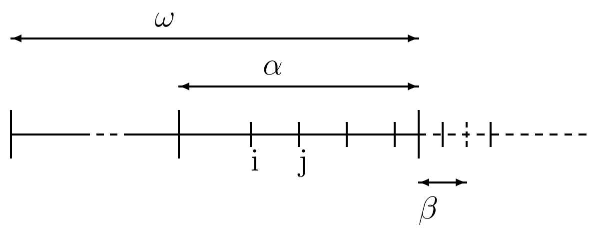
[Rub98] A. Rubinstein, 1998. *Modelling Bounded Rationality* , MIT Press.

[Sor86] S. Sorin, 1986. *On Repeated Games with Complete Information* , Mathematics of Operations Research, vol. 11 pp 147 − 161 .

Appendices

Proof of proposition 1

Let $n \in \mathbb{N}^*$ and $M^1 = [Q, q_*, f, g] \in \mathcal{M}^1(n)$ be an automaton for player i.
Let $\omega \in H_N$ such that $\sigma(M^1) \rightsquigarrow \omega$.
Let $\alpha \in \mathcal{C}^1_n$, $\alpha \triangleleft\!\rrbracket\, \omega$.
Setting $q_0 = q_*$ consider $q : \begin{pmatrix} [\![1, n+1]\!] & \to & Q^1 \\ t & \mapsto & f\left(q_{t-1}, \omega^2_{t-1}\right) \end{pmatrix}$; then $|Q^1| = n$ and thus q can not be injective : let $i < j$ be the first stages among $[\![1, n+1]\!]$ such that $q_i = q_j$, and set $T = j - i$.



We have $\omega_{N-n+i} = (g^1(q_i), \phi(g^1(q_i))) = (g^1(q_j), \phi(g^1(q_j))) = \omega_{N-n+j}$.
Thus by induction $\forall t \in [\![0, p-1]\!]$, $\omega_{N-n+i+t} = \omega_{N-n+j+t}$.
Let $s = T \texttt{ mod } p$ et $\beta = (\omega_{i+r+1}, \ldots, \omega_{i+p}, \omega_i, \ldots, \omega_{i+r})$.
We have $\sigma(M^1) \rightsquigarrow \omega\beta\beta\beta\ldots$, which completes the proof.

□

Lemma 4

The proof of lemma 1 relies on the following lemma :

Lemma 4 *Let $f : (\mathbb{N} \to X)$ be a periodic function of period n, and $g : (\mathbb{N} \to X)$ one of period k, such that $\forall t \in [\![0, n+k-1]\!],\ f(t) = g(t)$. Then $\forall t \in \mathbb{N},\ f(t) = g(t)$.*

Proof Let f and g be two such functions, and assume without loss of generality that $1 \leq k \leq n$. Let us show that $\forall t \in [\![0, 2(n+k)-1]\!],\ f(t) = g(t)$ first.
If k divides n, then g is also periodic of period n and the assertion is trivial.
Otherwise, let $r = n \texttt{ mod } k \in [\![1, k-1]\!]$ the reminder of the division of n by k.
Let $t \in [\![0, 2(n+k)-1]\!]$.
If $t \leq n+k$, the result is true by hypothesis; otherwise, let $s = t-(n+k) \in [\![0, n+k-1]\!]$.
On one hand we have

$$\begin{array}{lcll}
f(t) & = & f(n+k+s) & \text{by definition of } s \\
& = & f((k+s) \texttt{ mod } n) & \text{since } f \text{ is periodic of period } n \\
& = & f((k \texttt{ mod } n) + (s \texttt{ mod } n)) & \text{sine } (k+s) \texttt{ mod } n = ((k \texttt{ mod } n) + (s \texttt{ mod } n)) \texttt{ mod } n \\
& = & f(k + (s \texttt{ mod } n)) & \text{since } k \texttt{ mod } n = k \\
& = & g(k + s \texttt{ mod } n) & \text{by hypothesis, since } k + s \texttt{ mod } n \leq k+n-1 \\
& = & g(s \texttt{ mod } n) & \text{since } g \text{ is periodic of period } k
\end{array}$$

On the other hand we have

$$\begin{array}{lcll}
g(t) & = & g(n+k+s) & \text{by definition of } s \\
& = & g((n+s) \texttt{ mod } k) & \text{since } g \text{ is periodic of period } k \\
& = & g((n \texttt{ mod } k) + (s \texttt{ mod } k)) & \text{since } (n+s) \texttt{ mod } k = ((n \texttt{ mod } k) + (s \texttt{ mod } k)) \texttt{ mod } k \\
& = & g(n + (s \texttt{ mod } k)) & \text{since } g \text{ is periodic of period } k \\
& = & f(n + (s \texttt{ mod } k)) & \text{by hypothesis, since } n + (s \texttt{ mod } k) \leq n+k-1 \\
& = & f(s \texttt{ mod } k) & \text{since } f \text{ is periodic of period } n \\
& = & g(s \texttt{ mod } k) & \text{by hypothesis since } (s \texttt{ mod } k) \leq k-1 \leq n+k-1
\end{array}$$

And for $j \in [\![0, n+k-1]\!]$ we have

$$\begin{array}{rcll} g(j \texttt{ mod } k) & = & g(j) & \text{since } g \text{ is periodic of period } k \\ & = & f(j) & \text{by hypothesis, since } j \leq n+k-1 \\ & = & f(j \texttt{ mod } n) & \text{since } f \text{ is periodic of period } n \\ & = & g(j \texttt{ mod } n) & \text{by hypothesis, since } j \texttt{ mod } n \leq n-1 \leq n+k-1 \end{array}$$

Thus $g(t) = f(t)$, that is

$$\forall t \in [\![0, 2(n+k)-1]\!],\ f(t) = g(t)$$

To complete the proof of the lemma, observe that f and g are also periodic of periods $2n$ and $2k$ respectively, so that applying the same proof leads to $\forall t \in [\![0, 4(n+k)-1]\!],\ f(t) = g(t)$.
Hence by a simple induction

$$\forall t \in \mathbb{N}^*,\ f(t) = g(t)$$

□

Proof of lemma 1

Let $n \in \mathbb{N}$ and $\alpha^1 \in (A^1)^n$.
Let $\sigma^1 \in C_k(\omega)$ be such that $\sigma^1 \rightsquigarrow \omega\underline{\alpha^1}$.
Let $\mu^1 \in C_k(\omega)$; let us show that $\mu^1 \rightsquigarrow \omega\underline{\alpha^1}$.
Note that $(\omega_{k-2T+1}, \omega_k)$ is a best-reply path, and thus is in correspondance for player one.
Since σ^1 is T-cyclic, let $t_0 \in [\![1, T]\!]$ and $\alpha \in H_{t_0}$ tail of ω such that $\forall n \in \mathbb{N}$, $\sigma^1 \rightsquigarrow \omega\underbrace{\alpha\ldots\alpha}_{n \text{ times}}$; let $n_0 = T - t_0$. Note that α is a best-reply path.
Define conversely $t_1 \in [\![1, T]\!]$ and $\beta \in H_{t_1}$ for strategy μ^1, and let $n_1 = T - t_1$.
Assume without loss of generality that $n_1 \leq n_0$, and define

$$c^1 : \left(\begin{array}{ccc} \mathbb{N} & \rightarrow & A^1 \\ t & \mapsto & \sigma^1\left(\left[\left(\omega_{|k-2T+n_0}\right) \alpha\alpha\ldots \right]_{|k-2T+n_0+t} \right) \end{array} \right)$$

Then c^1 is a t_0-periodic function for any $t \in \mathbb{N}$. Define

$$d^1 : \left(\begin{array}{ccc} \mathbb{N} & \rightarrow & A^1 \\ t & \mapsto & \mu^1\left(\left[\left(\omega_{|k-2T+\mathbf{n_1}}\right) \beta\beta\ldots \right]_{|k-2T+\mathbf{n_0}+t} \right) \end{array} \right)$$

Then d^1 is t_1-periodic (even if $n_1 < n_0$).
Moreover $\forall t \in [\![0, 2T-n_0-1]\!]$, $c^1_t = \omega^1_{2T-n_0+t} = d^1_t$. It follows from lemma 4 that $\forall p \in \mathbb{N}, f(p) = g(p)$, that is $\mu^1 \rightsquigarrow \omega\underline{\alpha^1}$.

□