// In Preparation //

Completely Uncoupled Dynamics and Nash Equilibrium*

Babichenko Yakov†

March 19, 2009

Abstract

Completely uncoupled dynamic is a repited play of a game, when in every given time the action of every player depends only on his own payoffs in the past. In this paper we try to formulate the minimal set of necessary conditions that guarantee a convergence to a Nash equilibrium in completely uncoupled model.

The main results are:

1. The convergence to a Nash equilibrium cannot be guaranteed with finite memory strategies, in a generic game.

2. A convergence to an ε-Nash equilibrium almost all the time can be guaranteed with finite memory strategies, in a generic game.

1 Introduction

Uncoupled dynamics is a process of a repeated play of a one shot game, when strategy of every player does not depend on payoff functions of other players. The problem of convergence of uncoupled dynamics to equilibrium

*The author wishes to thenk his supervisor, Sergiu Hart, for his support and guidence, and Itai Arieli and Ron Perets for useful comments and discussions.

†Center for the Study of Rationality and Institute of Mathematics, The Hebrew University of Jerusalem, Feldman Building, Givat Ram, 91904 Jerusalem, Israel. E-mail: yak@math.huji.ac.il.

is well studied and widely represented in the literature. There are several reasonable uncoupled dynamics that lead the play to Nash equilibrium, for example the hypothesis testing of Foster and Young [4], or the public learning process introduced by Foster and Kakade [2].

Another strategy which can lead to Nash equilibrium is the exhaustive search. The idea is to go through all the possible actions in some order (deterministic exhaustive search) or randomly (probabilistic exhaustive search), until players get to some situation where they all are sure that they play a Nash equilibrium, so they stay to play it forever. In [8] Hart and Mas Colell showed among other results, that convergence to Nash equilibrium (pure or ε-equilibrium) can be guaranteed by using finite memory exhaustive search strategies.

Negative results (i.e., impossibility of convergence to Nash equilibrium) for uncoupled dynamics have been studied by Hart and Mas Colell in [7] for continuous time dynamics and in [8] for discrete time (see also Foster and Young [3], and Young [9]).

Completely uncoupled dynamic[1] is a repeated play of a game, when in every given time the action of every player depends only on his own payoffs in the past. The assumption is that a player knows neither his payoff function nor the actions played in the history. Actually he doesn't even know how many players participate in the game. There are several completely uncoupled dynamics that lead to Nash equilibrium:

Arslan, Marden,Shamma and Young [1] focused in acyclic games, and showed several strategies that lead to Nash equilibrium in this class of games.

The regret testing strategies of Foster and Young [5] are finite memory strategies that guarantee a convergence to ε-Nash equilibrium in $1 - \varepsilon$ of the time, in every two-players game. Germano and Lugosi [6] generalized the regret testing strategies to multy-players games, but added the condition that the game has to be generic. We can change the regret testing strategies to search for ε-Nash equilibrium for some constant time, then reduce ε and search for ε-Nash equilibrium for some (larger) constant time. If ε tends to 0, then we will get that the play converges to Nash equilibrium almost all the time, but then the strategies are infinite memory strategies.

Young's iteractive trail and error learning [10], are finite memory strategies, that guarantee a convergence to pure Nash equilibrium for every generic

[1] The concept called in the lierature also radicaly uncoupled strategies, or payoff based strategies.

multi-player game in $1-\varepsilon$ of the time.

We can see that there is a gap between the case of uncoupled dynamics and completely uncoupled dynamics, which can be formulated as the following questions:

1. In the uncoupled case, the convergence to Nash equilibrium can be guaranteed by finite memory strategies, whereas in the completely uncoupled case only infinite memory strategies that establish a convergence to Nash equilibrium, are known. The question is, whether it can be done with finite memory.

2. In the uncoupled case, the convergence to ε-Nash equilibrium can be guaranteed for general game, whereas in the completely uncoupled case only strategies that guarantee convergence for generic game are known. The question is whether it can be done for general game.

3. Convergence almost all the time to ε-Nash equilibrium can be guaranteed by finite memory strategies in the uncoupled case, whereas in the completely uncoupled case only strategies that guarantee the convergence in $1-\varepsilon$ of the time are known. The question is, whether we can improve the convergence to almost all the time?

In this paper we are going to answer these questions. Moreover, the goal is to formulate the minimal set of necessary conditions that guarantee a convergence to a Nash equilibrium in completely uncoupled model. Trying to answer the questions 1-3 formulated above, we show that:

1. The convergence to a Nash equilibrium cannot be guaranteed with finite memory strategies in a generic game (Corollary 2).

2. The assumption of the generic game is necessary for uncoupled dynamics to converge to an equilibrium (Theorem 9).

3. A convergence almost all the time to an ε-Nash equilibrium can be guaranteed with finite memory strategies in a generic game (Corollary 12).

The paper organized as follows. In Section 2 we introduce the model and notations. In Section 3 we study the question of convergence to pure Nash equilibria, and in Section 4 – to the ε-Nash equilibria, when the proofs of the theorems in both sections 3 and 4 postponed to section 5 for more transparent presentation. All the main proofs presented in Section 5.

2 The Model

In this section we mainly describe our notations and define the objects and concepts which will used in the paper. Part of them are standard, but we recall them for convenience of a reader.

2.1 The Game

A basic static (one-shot) game Γ is given in strategic form as follows. There are $n \geq 2$ players, denoted by $i = 1, 2, ..., n$. $N = \{1, 2, ..., n\}$ is the set of all the players. C is a countable set of all the possible actions of the players. Each player i has a finite set of pure actions $A^i = \{a_1^i, a_2^i, ..., a_{m^i}^i\} \subset C$ where $|A^i| = m^i$; let $A := A^1 \times A^2 \times ... \times A^n$ be the set of action profiles. Let $\mathcal{B}$, be the set of all actions for a single player, i.e., $\mathcal{B}$ is the set of all the finite subsets of C. Let $\mathcal{A}$ be the set of all the action profiles sets, i.e. $\mathcal{A} = \mathcal{B}^n$.

The payoff function (or utility function) of player i is a real valued function $u^i : A \to \mathbb{R}$. In this paper we assume that all the utility functions are bounded from above by some constant M big enough; i.e., $|u^i(a)| < M$ for every $i \in N$, and every $a \in A$. The set of mixed (or randomized) actions of player i is probability simplex over A^i

$$\Delta(A^i) = \{(p_j^i)_{j=1,...,m^i} | \Sigma_{j=1}^{m^i} p_j^i = 1 \text{ and } p_j^i \geq 0 \text{ for } j = 1, ..., m^i\}$$

The payoff functions u^i are multilinearly extended from A to $\Delta(A)$:

$$u^i : \Delta(A^1) \times \Delta(A^2) \times ... \times \Delta(A^n) \to \mathbb{R}$$

Let U_A^i be the set of all the payoff functions of player i (bounded by M). Let $U_A = U_A^1 \times U_A^2 \times ... \times U_A^n$ be the set of all the payoff functions of all the players when the action profile set is A. Let $\mathcal{U}$ be the set of all the games with every possible action profile set A

$$\mathcal{U} = \bigcup_{A \in \mathcal{A}} U_A$$

We denote all the games with n players by $\mathcal{U}_n$.

The actions of all the players except player i is $a^{-i} = (a^1, ..., a^{i-1}, a^{i+1}, ..., a^n)$, and the set of actions of all the players except player i is $A^{-i} = A^1 \times ... \times A^{i-1} \times A^{i+1} \times ... \times A^n$. Let $U_{A^{-i}}^{-i}$ be the set of all the payoff functions of all the players except player i.

An action $a^i \in A^i$ will be called a *best reply to* a^{-i} if $u^i(a^i, a^{-i}) \geq u^i(\overline{a}^i, a^{-i})$ for every $\overline{a}^i \in A^i$. A *pure Nash equilibrium* is an action profile $a = (a^1, a^2, ..., a^n) \in A$, such that a^i is a best reply to a^{-i} for all i.

2.2 The Dynamic Setup

The dynamic setup consists of the repeated play, at discrete times $t = 1, 2, ...$, of the static game U_A. Let $a^i(t) \in A^i$ denotes the action of player i at time t, and put $a(t) = (a^1(t), a^2(t), ..., a^n(t)) \in A$ for the combination of actions at time t.

We assume that at the end of time t each player i observes the action that he played $a^i(t)$ and his own payoff $u^i(a(t))$. At the time t player i knows his previous acts and payoffs which will be denoted by $h^i(t) = ((a^i(t'))_{t'=1}^t, (u^i(a(t'))_{t'=1}^t)$. Let $\mathcal{PH}_{t,A^i}$ be the set of all the histories of a player with actions set A^i at time t.

The *history of play* in a game with action profile set A is $H_A(t) = (a(1), a(2), ..., a(t))$, where $a(t') \in A$ for every $t' \leq t$. Let $\mathcal{H}_{t,A}$ be the set of all the histories of play of t steps, and $\mathcal{H}_A^* := \bigcup_{t=0}^{\infty} \mathcal{H}_{t,A}$.

Let $PNE \subset A$ be the set of all the pure Nash equilibria of the game U_A. We will say that *the play almost surely converges to* PNE if

$$P(\lim_{t \to \infty} \frac{\#\{t' | t' \leq t, a(t') \in PNE\}}{t}) = 1$$

i.e., the frequency of times when a pure Nash equilibrium was played, converges to 1 when $t \to \infty$, with probability 1.

2.3 Strategy Mappings

A *completely uncoupled strategy of a player with actions set* A^i, f_{A^i}, is a sequence of functions $(f_1, f_2, ..., f_t, ...)$, where $f_t : \mathcal{PH}_{t-1,A^i} \to \Delta(A^i)$. Denote by F_{A^i} the set of all the completely uncoupled strategies of player i with actions set A^i. The set of all the completely uncoupled strategies, for all the actions sets is $\mathcal{F}$. A mapping $\varphi : \mathcal{B} \to \mathcal{F}$, that assigns a completely uncoupled strategy $\varphi(A^i) = f_{A^i} \in F_{A^i}$ for every cations set $A^i \in \mathcal{B}$, will be called *completely uncoupled strategy mapping*. For every given strategy mapping φ, in a game with action profile set $A = A^1 \times A^2 \times ... \times A^n$, the strategies of the players will be $(f_{A^1}, f_{A^2}, ..., f_{A^n})$. We denote by $f = (f_{A^1}, f_{A^2}, ..., f_{A^n})$ the *strategy profile*. The strategy profile defines a probabilistic play of the game. A strategy profile *leads to Pure Nash equilibrium* if the play almost surely converges to PNE, in those games where such an equilibrium exists.

A strategy f is called a *finite memory strategy* if it can be implemented by a finite automaton.

A history $H(t)$ will be called *realizable* by a strategy profile f if after t steps of play, according to the strategy profile f, there probability that the history will be $H(t)$ is positive.

2.4 Genericity

For every game u with n players and set of actions A, we can consider u^i as an element of $\mathbb{R}^{|A|}$, and u - as an element of $\mathbb{R}^{n|A|}$. Therefore we can define Lebesgue measure $\lambda(\Omega)$ of games set Ω as a measure in $\mathbb{R}^{n|A|2}$. In the same way we define the measure of a set $\Omega \subset U^i$ or $\Omega \subset U^{-i}$.

We will say that a certain property *is valid in every generic game with action profile set* A, if the property holds for all games with action profile set A except a subset of games with measure 0. We will say that a certain property is valid *in every generic game* if for every $A \in \mathcal{A}$ the property is valid in every generic game with action profile set A.

3 Pure Nash equilibrium

We will mainly concentrate on the case of generic games in the next paragraph. The case of not generic games will be considered separately below.

3.1 Generic Games

In generic games we are looking for strategies that will guarantee a convergence to a pure Nash equilibrium in *almost* very game.

The considerations will be divided in two cases: finite and infinite memory strategies.

[2]Below measure will be understood as Lebesgue measure.

3.1.1 Finite Memory Strategies

The following negative result shows that using finite memory strategies the convergence of play to a pure Nash equilibrium cannot be guaranteed.

Theorem 1 *Let $A = A^1 \times A^2 \times ... \times A^n$ be an action profile set such that $A^1 = A^2$. Then there is no completely uncoupled mapping into finite memory strategies leading to a pure Nash equilibrium in every generic game with action profile set A, and in every generic game with action profile set A^{-1}.*

$>$From this theorem immediately follows

Corollary 2 *There is no completely uncoupled mapping into finite memory strategies that leads to a pure Nash equilibrium in every generic game.*

If one would like to formulate a positive statement about convergence to pure Nash equilibrium in the case of finite memory strategies, there arises a problem to save (to remember) some observation from the past, since the payoffs are real numbers, and it is impossible to save a general real number in a finite memory. An approach to avoid this problem is to consider a class of games $DP_{A,\delta}$ where the payoffs of a player are different one from another by at least δ:

$$DP_{A,\delta} = \{u \in \mathcal{U} : |u^i(a) - u^i(a')| \geq \delta \text{ for every } i \in N \text{ and every } a \neq a', \text{ with } a, a' \in A\}$$

We denote $DP_\delta = \underset{A \in \mathcal{A}}{\cup} DP_{A,\delta}$.

In the classes of games $DP_{A,\delta}$ and DP_δ the player can save just the first $-\log \delta$ digits after the decimal point, and still he could distinguish between two different observations. The following Lemma shows that for δ small enough $DP_{A,\delta}$ is very close to U_A.

Lemma 3 *For every action profile set A, and for every $\varepsilon > 0$ there exists $\delta > 0$ such that $\lambda((DP_{A,\delta})^C) < \varepsilon$.*

Proof. By definition:

$$(DP_{A,\delta})^C = \{u \in \mathcal{U} : \exists i \in N, \exists a \neq a', a, a' \in A \text{ such that } |u^i(a) - u^i(a')| < \delta\}$$

For every player $i \in N$, and for every pair of actions $a, a' \in A$,$a \neq a'$, let $E_{\delta,i,a,a'} = \{u \in \mathcal{U} : |u^i(a) - u^i(a')| < \delta\}$. The measure of the set $E_{\delta,i,a,a'}$ is

$\lambda(E_{\delta,i,a,a'}) = (2M)^{n|A|-1} \cdot 2\delta$, because all the payoffs of all the players except player i at the action a' could be any number in the segment $(-M, M)$, whilst the payoff of player i at the action a' is in the segment $(u^i(a) - \delta, u^i(a) + \delta)$.

Clearly

$$(DP_{A,\delta})^C = \bigcup_{i \in N,\ a,a' \in A\ a \neq a'} E_{\delta,i,a,a'}$$

therefore

$$\lambda((DP_{A,\delta})^C) \leq \sum_{i \in N,\ a,a' \in A\ a \neq a'} \lambda(E_{\delta,i,a,a'}) = n \cdot \binom{|A|}{2} \cdot (2M)^{n|A|-1} \cdot 2\delta$$

so for every $\varepsilon > 0$ there exists small enough $\delta > 0$ such that

$$n \cdot \binom{|A|}{2} \cdot (2M)^{n|A|-1} \cdot 2\delta < \varepsilon \tag{1}$$

and for this δ: $\lambda((DP_{A,\delta})^C) < \varepsilon$ as required. ■

Let $DA \subset \mathcal{U}$ be the set of all the action profile sets in which all the players have a different actions sets; i.e., $A^i \neq A^j$ for all $i \neq j$.

In Theorem 1, we considered games with action profile set $A = A^1 \times A^2 \times \ldots \times A^n$, such that $A^1 = A^2$. Clearly the Theorem remains valid for action profile set A with any two equal actions sets ($A^i = A^j$ for $i \neq j$). On the other hand, the following Theorem claims that for every game with different actions sets (DA) in which the players could remember their payoffs in a finite memory (DP_δ), the convergence to a pure Nash equilibrium, could be guaranteed by finite memory strategies.

Theorem 4 *For every $\delta > 0$ there exists a completely uncoupled mapping into finite memory strategies which leads to a Pure Nash equilibrium in every game $\Gamma \in DA \cap DP_\delta$ where such an equilibrium exists.*

Corollary 5 *Assume that the number of players is bounded by P, and the number of actions of every player is bounded by T. Then for every $\varepsilon > 0$ there exists a completely uncoupled mapping into finite memory strategies that leads to a Pure Nash equilibrium in every game with different actions sets, except a set of games with measure smaller than ε.*

Proof of the Corollary. Let $\widetilde{A}$ be an action profile set of P players, where every player has T actions. For every $\varepsilon > 0$ there exists $\widetilde{\delta} > 0$ such that $\lambda((DP_{\widetilde{A},\widetilde{\delta}})^C) < \varepsilon$, by Lemma 3. For every action profile set A, with number of players less then P and number of actions less then T one can see by inequality (1) that the same $\widetilde{\delta} > 0$ guaranties that $\lambda((DP_{A,\widetilde{\delta}})^C) < \varepsilon$. For every actions profile set A, by Theorem 4 there exists a completely uncoupled strategy mapping that guarantees a convergence to Nash equilibrium in every game $\Gamma \in DA \cap DP_{A,\widetilde{\delta}}$. Therefore the measure of the games for which the convergence to a Nash equilibrium is not guarantied is $\lambda((DP_{A,\widetilde{\delta}})^C) < \varepsilon$. ■

Complete uncoupledness with additional information There exists a basic information of a strategy (a player) before the game starts- the domain of the strategy mapping. Above the basic information of the players was their action set only. Let us consider the case when there is some additional information for every player. It makes changes in the model described above: we now allow the strategies of the players dependent not only on the actions set.

Let K be the set of all the possible values of information. For example, if the information is the index number of the player, then $K = \mathbb{N}$. Let $\alpha = (\alpha^1, \alpha^2, ..., \alpha^n) : \mathcal{U}_n \to K^n$ be the *information function*, when $\alpha^i(u)$ is the information that of player i. In the example above $\alpha^i(u) = i$. An information is called *uncoupled information*, if the information is not about the payoff functions of the other players, or more formally, for every two games $u_1 = (u^i, u^{-i})$ and $u_2 = (u^i, \overline{u}^{-i})$ holds $\alpha^i(u_1) = \alpha^i(u_2)$.

Till now we had $\varphi : \mathcal{B} \to \mathcal{F}$ a strategy mapping. Now we want the strategy mapping to be from $\mathcal{B} \times K$ to $\mathcal{F}$. Let $\varphi : \mathcal{B} \times K \to \mathcal{F}$ be a *completely uncoupled strategy mapping with additional information* α. Given a strategy mapping φ, for every game u with action profile set $A = A^1 \times A^2 \times ... \times A^n$, the strategies of the players will be $(\varphi(A^1, \alpha^1(u)), \varphi(A^2, \alpha^2(u)), ..., \varphi(A^n, \alpha^n(u)))$.

In the following Theorem we show two examples of additional information (one may say a reasonable information), that guarantee the convergence to a pure Nash equilibrium. Like in Theorem 4 we restrict the games to the class DP_δ, where the players could save their payoffs in finite memory.

Theorem 6 *For every* $\delta > 0$*, if*
1) for every player $i \in N$*,* α^i *is the index of the player, or*
2) for every player $i \in N$*,* α^i *is the total number of the players in the game, then there exists a completely uncoupled mapping with additional information*

α into finite memory strategies that leads to a Pure Nash equilibrium in every game $\Gamma \in DP_\delta$, where such an equilibrium exists.

Similar to Theorem 4, we can formulate the following corollary

Corollary 7 *Assume that the number of players is bounded by P, and the number of actions of every player is bounded by T. Then if*
1) for every player $i \in N$, α^i is the index of the player, or
2) for every player $i \in N$, α^i is the total number of the players in the game, then for every $\varepsilon > 0$ there exists a completely uncoupled mapping with additional information α into finite memory strategies that leads to a Pure Nash equilibrium in every game where such an equilibrium exists except a set of games with measure smaller than ε.

And the proof of it is also similar to the proof of Corollary 5.

3.1.2 Infinite memory strategies

Unlike the finite memory strategies, in the infinite memory case, the convergence to a pure Nash equilibrium can be guaranteed generically (i.e., for almost all the games).

Theorem 8 *There exists a strategy mapping $\varphi = \{f_m\}_{m=2}^{\infty}$ into infinite memory strategies that leads to a Pure Nash equilibrium in almost every game Γ, where such an equilibrium exists.*

3.2 Not generic games

The case of not generic games is less interesting because of the following strong negative statement.

Theorem 9 *There are no completely uncoupled strategy mapping with additional uncoupled information, that leads to a pure Nash equilibrium, in every game with more then 2 players, where such an equilibrium exists.*

The Theorem claims, not only that there is no strategies that leads to a pure Nash equilibrium in every game, but also if the players has some uncoupled additional information about the game, still such a strategies do not exist.

Proof. Consider the following two 3-players games:

Γ_1 :

	a_1^2	a_2^2
a_1^1	$1,1,1$	$1,1,1$
a_2^1	$1,1,1$	$1,1,1$

a_1^3

	a_1^2	a_2^2
a_1^1	$1,0,1$	$0,1,1$
a_2^1	$0,1,1$	$1,0,1$

a_2^3

Γ_2 :

	a_1^2	a_2^2
a_1^1	$1,0,1$	$0,1,1$
a_2^1	$0,1,1$	$1,0,1$

a_1^3

	a_1^2	a_2^2
a_1^1	$1,1,1$	$1,1,1$
a_2^1	$1,1,1$	$1,1,1$

a_2^3

The Pure Nash equilibria of Γ_1 are $\{(i,j,1)\}_{i,j=1}^2$. The Pure Nash equilibria of Γ_2 are $\{(i,j,2)\}_{i,j=1}^2$.

The strategy of player 3, in both games is the same strategy, and the histories of player 3 in both games are the same histories. So for one of the actions 1 or 2, player 3 does not play it with frequency that converges to 1 with probability 1. If this action is action i $(i = 1,2)$, then at the game Γ_i the strategies will not lead to a pure Nash equilibrium.

For different number of actions or different number of players, one can easily construct a similar example in which the leading to a pure Nash equilibrium cannot be guaranteed. ■

4 ε-Nash Equilibrium

In this section we assume that the number of players is bounded by P, and the number of actions of every player is bounded by T.

Let $x = (x^1, x^2, ..., x^n)$, $x^i \in \Delta(A^i)$ be a mixed actions profile. x^i is an *ε-best reply to x^{-i}* if $u^i(x^i, x^{-i}) \geq u^i(y^i, x^{-i}) - \varepsilon$ for every $y^i \in \Delta(A^i)$. $x = (x^1, x^2, ..., x^n)$ is an *ε-Nash equilibrium,* if for every $i \in N$, x^i is an ε-best reply to x^{-i}.

Given a strategy profile f we will say that the *play converges to an ε-Nash equilibrium* if the mixed actions of the players converge to ε-Nash equilibrium almost for every history[3]; i.e., for almost every history, for all player $i \in N$ the limit $\lim_{t\to\infty} f^i(H(t-1))$ exists and is equal to x^i, and $x = (x^1, x^2, ..., x^n)$ is an ε-Nash equilibrium.

[3] This kind of convergence called also *period by period behavior convergnce.*

Exactly by the same proof as for Theorem 1, we can show the following version of Theorem 1 for the case of pure Nash equilibrium:

Claim 10 *Let $A = A^1 \times A^2 \times ... \times A^n$ be an action profile set such that $A^1 = A^2$, then there is no completely uncoupled mapping into finite memory strategies that leads to a pure Nash equilibrium in every generic game $\Gamma \in DP_{A,\delta}$, and in every generic game $\Gamma \in DP_{A^{-1},\delta}$.*

Unlike the case of pure Nash equilibrium, the following theorem claims that the convergence to an ε-Nash equilibrium can be guaranteed by finite memory strategies on the class of games DP_δ.

Theorem 11 *For every $\varepsilon > 0$, and $\delta > 0$ there exists a completely uncoupled mapping into finite memory strategies that leads to a ε-Nash equilibrium in every game $\Gamma \in DP_\delta$.*

As in the other theorems about the class DP_δ we have the following corollary.

Corollary 12 *For every $\varepsilon > 0$ there exists a completely uncoupled mapping into finite memory strategies that leads to an ε-Nash equilibrium in every game except a set of games with measure smaller than ε.*

Finally similar to the case of pure Nash equilibrium, the convergence to ε-Nash equilibrium can be guaranteed by infinite memory strategies in generic games.

Theorem 13 (Analog of ε-Nash equilibrium to Theorem 9) *For every $\varepsilon > 0$ there exists a completely uncoupled mapping into finite memory strategies that leads to an ε-Nash equilibrium in every generic game.*

Note that this theorem is not straightforwardly implied by Theorem 11 (or Corollary 12), because it states that there exist strategies that leads to an ε-Nash equilibrium not only for games in the class DP_δ, but for almost all the games.

5 Proofs

Proof of Theorem 1. Assume, by contradiction, that such a strategy mapping φ exists. Note that $A^{-1} = A^{-2}$, therefore φ leads to a pure Nash equilibrium in every generic game with action profile sets A, A^{-1} and A^{-2}. By using the fact that φ leads to pure Nash equilibrium in generic game with action profile set A^{-1} and A^{-2}, we will prove that there exists a set of payoff functions $P \subset U_A$ with a positive measure such that

(i) for every $\Gamma \in P$, Γ has a pure Nash equilibrium

(ii) the strategies $(f_{A^i})_{i=1}^n$ do not lead to it.

It will show that the property of leading to Nash equilibrium doesn't hold for generic games.

Let $SPN_{A^{-1}} \subset U_{A^{-1}}$ be the set of all the payoffs which has a single pure Nash equilibrium, and the payoffs are bounded by $M-5$, where M is the bound for the payoffs in all the games. Then $\lambda(SPN_{A^{-1}}) > 0$, see Lemma 14. For every $v \in SPN_{A^{-1}}$ let $b(v)$ be the single pure Nash equilibrium. The strategies $(f_{m^i})_{i=2}^n$ of players $2, 3, ..., n$ lead to $b(v)$ for almost every game v. Let $S \subset SPN_{A^{-1}}$ be the set of the games for which $(f_{m^i})_{i=2}^n$ leads to $b(v)$, then $\lambda(S) > 0$. By Lemma 15 there exists t and an history $H_{A^{-1}}(t) \in \mathcal{H}^*_{A^{-1}}$, realizable by $(f_{m^i})_{i=2}^n$, such that from the time $t+1$ and on, the players play $b(v)$ with probability 1. For every $H_{A^{-1}}(t) = H_{A^{-1}} \in \mathcal{H}^*_{A^{-1}}$ let T_H be the subset of all the games $v \in U_{A^{-1}}$ such that

- $H_{A^{-1}}$ is realizable by $(f_{m^i})_{i=2}^n$
- if $H_{A^{-1}} = H_{A^{-1}}(t)$ is played, then from the time $t+1$ and on the players play some action $b \in A^{-1}$ with probability 1.

Then

$$\bigcup_{H_{A^{-1}} \in \mathcal{H}^*_{A^{-1}}} T_{H_{A^{-1}}} \supset S$$

There is a countable number of histories $H_{A^{-1}}$. S has a positive measure, therefore there exists $\overline{H}_{A^{-1}} = \overline{H}_{A^{-1}}(\overline{t})$ such that $T_{\overline{H}_{A^{-1}}}$ has a positive measure. Denote it by $R := T_{\overline{H}_{A^{-1}}}$. The action that played from the time $\overline{t}$ and on is denoted by $(\overline{a^2}, \overline{a^3}, ..., \overline{a^n})$.

One should note that $A^1 = A^2$, so $f_{A^1} = f_{A^2}$; i.e., players 1 and 2 have the same strategy. Therefore by the same reasons for the action profile set A^{-2} for the same subset of $\mathbb{R}^{(n-1)|A|}$: $R := T_{\overline{H}_{A^{-2}}} \subset U^{-2}$ holds

- $\overline{H}_{A^{-2}}$ is realizable by $(f_{m^i})_{i=1, i\neq 2}^n$

- if $\overline{H}_{A^{-2}}$ is played, then from the time $t+1$ and on the players play the action $(\overline{a^1}, \overline{a^3}, \overline{a^4}, ..., \overline{a^n}) \in A^{-2}$ with probability 1.

Let us introduce a simplifying notation. For every $u \in U_A$ and for every subset of actions $B \subset A$, let $u|_B$ be the payoff function, defined just on the subset B.

Here we define our $P \subset U_A$ as the set of all the games with payoff function $u = (u^1, u^2, ..., u^n)$ such that on the diagonal $a^1 = a^2$ the payoffs $u^{-1}|_{\{a \in A | a^1 = a^2\}}$ and $u^{-2}|_{\{a \in A | a^1 = a^2\}}$ are some payoffs of the subset R. Out of the diagonal we want that the payoffs of all the players to be better then on the diagonal. Furthermore, we want a_2^1 to be a dominant [4] action for player 1, and a_1^i be the dominant action of player $i \neq 1$. Put the following payoffs:

For player 1

$$u^1(a^1, a^2, ..., a^n) \in \begin{cases} [M-2, M-1] \text{ for } a^1 = a_2^1 \\ [M-4, M-3] \text{ for } a^1 \neq a_2^1 \end{cases} \text{ for } a^1 \neq a^2$$

For every player $i \neq 1$

$$u^i(a^1, a^2, ..., a^n) \in \begin{cases} [M-2, M-1] \text{ for } a^i = a_1^i \\ [M-4, M-3] \text{ for } a^i \neq a_1^i \end{cases} \text{ for } a^1 \neq a^2$$

Formally we define $P \subset U$ to be the set of all the payoff functions $u = (u^1, u^2, ..., u^n) \in U$ such that:

(a) $u^{-1}(a^2, a^3, ..., a^n), u^{-2}(a^1, a^3, a^4, ..., a^n) \in R$ for every action $a = (a^1, a^2, ..., a^n) \in A$ such that $a^1 = a^2$.

(b) $M - 4 < u^i(a) \leq M - 3$ for every action a such that $a^1 \neq a^2$, $i \neq 1$ and $a^i \neq a_1^i$.

(c) $M - 4 < u^i(a) \leq M - 3$ for every action a such that $a^1 \neq a^2$, $i = 1$ and $a^i \neq a_2^i$.

(d) $M - 2 < u^i(a) \leq M - 1$ for every action a such that $a^1 \neq a^2$, $i \neq 1$ and $a^i = a_1^i$.

(e) $M - 2 < u^i(a) \leq M - 1$ for every action a such that $a^1 \neq a^2$, $i = 1$ and $a^i = a_2^i$.

We will show that P satisfies the following:

1. $\lambda(P) > 0$.
2. Every game $u \in P$ has a single pure Nash equilibrium.

[4] Note that the action is not a dominant action in the game, but just dominant for actions out of the diagonal.

3. For every game $u \in P$ there is a positive probability that the strategies $f_{A^1}, f_{A^2}, ..., f_{A^n}$, will not lead to the pure Nash equilibrium.

which will complete the proof.

1. For every payoff function u that satisfies (a)-(e), conditions (b)-(e) restrict the payoffs out of the diagonal to be in some segment with length (or measure) 1. So

$$\lambda(P) = \underbrace{\lambda_{\mathbb{R}^{n|A^{-1}|}}(P|_{\{a \in A | a^1 = a^2\}})}_{\text{on the diagonal}} \cdot \underbrace{1^{n(|A|-|A^{-1}|)}}_{\text{out of the diagonal } a^1 \neq a^2} = \lambda_{\mathbb{R}^{n|A^{-1}|}}(P|_{\{a \in A | a^1 = a^2\}}) \tag{2}$$

Let $B := \{b = (u^1, u^2, u^{-\{1,2\}}) | (u^1, u^{-\{1,2\}}) \in R \text{ and } (u^2, u^{-\{1,2\}}) \in R\}$. One can see that $B = P|_{\{a \in A | a^1 = a^2\}}$. By Lemma 16 with $k = |A^{-1}|$, $l = (n-2)|A^{-1}|$, $C = R$ we have $\lambda(P|_{\{a \in A | a^1 = a^2\}}) > 0$. Therefore by (2) $\lambda(P) > 0$.

2. If the players play some action on the diagonal, then for players 1 and 2 it is better to play some action out of the diagonal. If the players play out of the diagonal, then the action

$$\begin{cases} a_1^i \text{ for } i \neq 1 \\ a_2^1 \text{ for } i = 1 \end{cases}$$

is the only best reply. So $(a_2^1, a_1^2, ..., a_1^n)$ is a single Pure Nash equilibrium.

3. The histories $\overline{H}_{A^{-1}}(\bar{t})$ and $\overline{H}_{A^{-2}}(\bar{t})$ are equal. Therefore players $3, 4, ..., n$ plays during all the history $\overline{H}_{A^{-1}}(\bar{t})$ the same actions as in the history $\overline{H}_{A^{-2}}(\bar{t})$. Therefore we can denote by $\overline{H}_A(\bar{t})$ the history in which players 1 and 2 play the same actions as in the histories $\overline{H}_{A^{-1}}(\bar{t})$ and $\overline{H}_{A^{-2}}(\bar{t})$ correspondingly, and players $3, 4, .., n$ play the same actions as in both these histories.

The history $\overline{H}_A(\bar{t})$ is realizable by the strategies $(f_{A^i})_{i=1}^n$ in every game $u \in P$, because we defined the payoffs u^i on the diagonal to be payoffs from the set R, so in every period $t = 1, 2, ..., \bar{t}$ there is a positive probability that all the players $i = 1, 2, ..., n$, will continue to play $a^i(t)$ and then they will stay on the diagonal ($a^1 = a^2$).

So the history $\overline{H}_A(\bar{t})$ will occur with a positive probability, as a result the action $(\overline{a^1}, \overline{a^2}, \overline{a^3}, ..., \overline{a^n})$ will be played with probability 1, at all the periods $\bar{t} + 1$ and forward. But $(\overline{a^1}, \overline{a^2}, \overline{a^3}, ..., \overline{a^n})$ is not a Nash equilibrium in the game u (because $\overline{a^1} = \overline{a^2}$), so the strategies do not lead to a pure Nash equilibrium. ■

Lemma 14 *For every $A \in \mathcal{A}$ the subset $SPN \subset U_A$ of the games with a single pure Nash equilibrium, has a positive measure.*

Proof. Consider the following n-players game:

-$u^i(a) = 1$ for every a such that $a^i = a^i_1$.

-$u^i(a) = 0$ for every a such that $a^i \neq a^i_1$.

The action $a^i = 1$ is a dominant strategy for every player i, so the game has a single pure Nash equilibrium $(1, 1, ..., 1)$. Also every $\frac{1}{3}$-perturbation of this game has the same single pure Nash equilibrium. Since environment of size $\frac{1}{3}$ of every game has a positive measure, we found a subset of games with a positive measure as required. ■

Lemma 15 *Let $f = (f^1, ..., f^n)$ be a strategy profile, where every f^i is a finite memory strategy, that guarantees almost sure convergence of the play to PNE in a game Γ that has a single pure Nash equilibrium $a = (a^1, ..., a^n)$. Then there exists a history $H(t) = (a(1), a(2), ..., a(t))$, realizable by f, such that from period $t + 1$ and on, the players play the Nash equilibrium a with probability 1.*

Proof. Let Λ^i be the set of all the possible memory states of player i; i.e., all the states of the strategy automaton. Let $\Lambda = \Lambda^1 \times \Lambda^2 \times ... \times \Lambda^n$.

The strategies $f^1, ..., f^n$ induce a Markov process on the finite Markov chain Λ. By Perron-Frobenius Theorem there is a stationary distribution π over the Markov chain states Λ. Denote by $\Omega \subset \Lambda$ the support of π. For every state $\omega \in \Omega$ the played action is a (with probability one), because otherwise there is some other action that played with frequency that does not converges to zero. For every $\omega \in \Omega$ the transition to the next state will be to some states of Ω. There exists a time t and a path $(\alpha_1, \alpha_2, ..., \alpha_t)$ $\alpha_i \in \Lambda$, which is realized with a positive probability, such that $\alpha_t \in \Omega$. Let $H(t)$ be the history of play on the path $(\alpha_1, \alpha_2, ..., \alpha_t)$, then $H(t)$ realizable by the strategy profile f, and from period $t + 1$ and on, the players play the Nash equilibrium a with probability 1. ■

Lemma 16 *For every $k, l \in \mathbb{N}$ and for every set $C \subset \mathbb{R}^{k+l}$ with a positive measure $\lambda(C) > 0$, the set*

$$B := \{b = (x, y, z) \in \mathbb{R}^k \times \mathbb{R}^k \times \mathbb{R}^l | (x, z) \in C \text{ and } (y, z) \in C\} \subset \mathbb{R}^{2k+l}$$

has a positive measure.

Proof. Let $\mathbf{1}_C, \mathbf{1}_B$ be the characteristic functions of C, B. $\mathbf{1}_B(x, y, z) = \mathbf{1}_C(x, z)\mathbf{1}_C(y, z)$ by the definition of B. For every $z \in \mathbb{R}^l$ let $g(z) := \lambda(\{x \in$

$\mathbb{R}^k | (x,z) \in C\})$. By Fubini Theorem

$$0 < \lambda(C) = \int_{\mathbb{R}^l} \left(\int_{\mathbb{R}^k} \mathbf{1}_C(x,z)dx \right) dz = \int_{\mathbb{R}^l} g(z)dz \tag{3}$$

also, by Fubini Theorem

$$\begin{aligned} \lambda(B) &= \int_{\mathbb{R}^l}(\int_{\mathbb{R}^k}(\int_{\mathbb{R}^k}\mathbf{1}_B(x,y,z)dx)dy)dz = \\ &= \int_{\mathbb{R}^l}(\int_{\mathbb{R}^k}(\int_{\mathbb{R}^k}\mathbf{1}_C(x,z)\mathbf{1}_C(y,z)dx)dy)dz = \\ &= \int_{\mathbb{R}^l}(\int_{\mathbb{R}^k}\mathbf{1}_C(x,z)dx)(\int_{\mathbb{R}^k}\mathbf{1}_C(y,z)dy)dz = \int_{\mathbb{R}^l}g(z)g(z)dz \end{aligned} \tag{4}$$

By (3) $\int_{\mathbb{R}^l} g(z)dz > 0$, therefore by (4) $\lambda(B) = \int_{\mathbb{R}^l} g^2(z)dz > 0$. ■

Proof of Theorem 4. The set C of all the possible actions for all the players is countable, therefore the set $\mathcal{B}$ of all the finite subsets of C is also countable. So there is an injective function $\gamma : \mathcal{B} \to \mathbb{N}$. So for every game $\Gamma \in DA$ the numbers $\gamma(A^1), \gamma(A^2), ..., \gamma(A^n)$ are different.

For every game $\Gamma \in DP_\delta$ the payoffs of every player are different at least in δ, so given a payoff $u^i(a)$ for player i, the player could ascribe importance just to the first $1 - \log_{10}(\delta)$ digits after the decimal point of the payoff. We denote it by $\lfloor u^i(a) \rfloor_\delta$. Then $\lfloor u^i(a) \rfloor_\delta \neq \lfloor u^i(a') \rfloor_\delta$for every $a, a' \in A$ $a \neq a'$. So when we say that a strategy of player i remembers the payoff $u^i(a)$, we actually mean that the strategy remembers $\lfloor u^i(a) \rfloor_\delta$, which could be saved in a finite memory.

Let φ be a mapping that assigns the strategy $f(\gamma(A^i))$ for every actions set $A^i \in \mathcal{B}$, where the strategy $f_{A^i}(l)$ for $l \in \mathbb{N}$ is defined below.

We start from formal descriptions of strategy $f_{A^i}(l)$, and we will explain later why the strategies $f(\gamma(A^1)), f(\gamma(A^2)), ..., f(\gamma(A^n))$ lead to a pure Nash equilibrium.

The strategy $f_{A^i}(l)$ is composed from five main steps.

Step 1 is called the **evaluation of index** step:

Substep 1.1: The player plays a^i_1, and remembers his payoff $u^i(a^1_1, a^2_1, ..., a^n_1)$, denote it for short $u^i(1)$.

The player moves to the next substep 1.2.

Substep 1.2.k.1, for $k \in \mathbb{N}$: If $l = k$ (i.e., $\gamma(A^i) = k$) the player plays a^i_1. Otherwise he plays a^i_2. Anyway he remembers the payoff.

Substep 1.2.k.2: If $l > k$ the player plays a^i_2. Otherwise he plays a^i_1. If his payoff is $u^i(1)$ he evaluates his index j by

$$j = \#\{k|k < l \text{ and the player gets at step 1.2.}k\text{.1 different payoff from } u^i(1)\}+1$$

and continues to step 2. Otherwise if his payoff is not $u^i(1)$ he continues to substep 1.2.$k+1$.1.

Step 2, is called the **recognition of action profile set** step:

Below we will show that the indexes of different players are different, so the player with index j will be called player j.

Substep 2.k for $k \neq j$:

Substep 2.k.1 :The player plays a^i_1. If his payoff is $u^i(1)$ he concludes that there are $k-1$ players and he continues to step 3. Otherwise he continues to substep 2.k.2.

Substep 2.k.l, for $l \geq 2$: The player plays a^i_1. If his payoff is $u^i(1)$ he concludes that player k has l actions and he continues to substep 2.$k+1$. Otherwise he continues to substep 2.k.$l+1$.

Substep 2.j:

The player plays his actions by the following order: $a^i_2, a^i_3, ..., a^i_{m^i}, a^i_1$ and then continues to substep 2.$j+1$.

Step 3, is called the **recognition of payoff function** step:

Below we will show that after step 2, the player knows the profile action set A.

The player goes through all the actions $a \in A$ by the lexicographic order (the order is defined by their indexes j). For every $a \in A$ the player plays his action a^i and remembers the payoff.

Step 4, is called the **finding a pure Nash equilibrium** step:

The player knows his payoff function from step 3.

Looking through all the actions $a = (a^i, a^{-i})$ by the lexicographic order the player plays a^i_1 if a^i is a best reply to a^{-i}, otherwise he plays a^1_2. If his payoff is $u^j(1)$ then he remembers a^i as the pure Nash equilibrium action and moves to step 5. Otherwise he goes to the next, by the lexicographic order, action.

Step 5 is called the **playing the pure Nash equilibrium** step, which is just the repeated play if his action in the pure Nash equilibrium.

Each substep can be implemented by a finite automata, so the strategies are finite automata strategies.

Let's explain why the strategies $f(\gamma(A^1)), f(\gamma(A^2)), ..., f(\gamma(A^n))$ lead to a pure Nash equilibrium.

At step 1: The players will go through all the natural numbers $k = 1, 2, ..., \max_{i \in N} \gamma(A^i)$. For every number k the players will know at step $1.2.k.1$ whether there exists a player i with $\gamma(A^i) = k$, or not. When they will get to $k = \max_{i \in N} \gamma(A^i)$, at the step $1.2.k.2$ they will know that there is no player i such that $\gamma(A^i) > k$. $\{\gamma(A^i)\}_{i=1}^n$ are different, so the indexes

$$j(i) = \#\{k \in N | \gamma(A^k) < \gamma(A^i)\} + 1$$

are also different. In addition $\{j(1), j(2), ..., j(n)\} = \{1, 2, ..., n\}$.

At step 2: First player i with index $j(i) = 1$ will play his actions $a_2^i, a_3^i, ..., a_{m^i}^i, a_1^i$. When he will finish this process, all the players will know it, because their payoff will be $u^i(1)$. Hence the players will know the number of actions of player 1. After that, the same will happen with player i' who's index is $j(i') = 2$, and so on, till player n. When it will be the turn of player $n+1$, the players will see that player $n+1$, has only one action, and it means that player $n+1$ does not exist. At the end of step 2, the players will know the total number of players, and the number of actions of every player.

At step 3: The players will play all the possible actions $a \in A$, by the lexicographic order, so they will know their utility function.

At step 4: The players will look through all the possible actions $a \in A$, until there will be an action $a \in A$ in which their payoffs will be $u^i(1)$. It means that all the players make a best reply action at the action profile a. Therefore a is a pure Nash equilibrium, and at step 5 the players will play it all the time.

Given a game $\Gamma \in DA$ with a pure Nash equilibrium, all the players will go through all the steps simultaneously, and eventually they will get to step 5, where they play a Nash equilibrium all the time, so the frequency of times that the players play a Nash equilibrium converges to 1. ■

Proof of Theorem 6. As in the proof of Theorem 4, the players can save every payoff in every game $\Gamma \in DP_\delta$ in a finite automata.

Let us prove that each one of the two conditions is sufficient:

Condition 1: If all the players have some different indexes from $\{1, 2, ..., n\}$, then by Theorem 4, the chain of steps:

"recognition of action profile set" $\rightarrow$ "recognition of payoff function" $\rightarrow$ "finding of pure Nash equilibrium" $\rightarrow$ "playing the pure Nash equilibrium"

guarantees a convergence to a pure Nash equilibrium.

By the assumption, the players know their index i, so this four steps will lead to a pure Nash equilibrium.

Condition 2: The strategy mapping of every player i can depend on n, so let us define the strategy of player i in the following way:

The player randomizes uniformly a natural number $1 \leq c^i \leq n$. Afterwards he plays the step "recognition of action profile set", as if his index is $j = c^i$. If at the end of this step he concludes that there is n players, he continues the steps chain "recognition of payoff function" $\rightarrow$ "finding of pure Nash equilibrium". Otherwise he randomizes again a number $1 \leq c^i \leq n$ and repeats the process. We call this strategy "*finding a pure Nash equilibrium with n players*". Finally when he found a pure Nash equilibrium, he plays it.

When each player will use this strategy, the following will happen:

If $(c^1, c^2, ..., c^n)$ is a permutation of $(1, 2, ..., n)$, then after the "recognition of action profile set" all the players will conclude that there are n players, and finally find a pure Nash equilibrium.

If at least two of the players have randomized the same number, let j be the smallest number such that $j \notin \{c^1, c^2, ..., c^n\}$. At the end of step 2 all the players will conclude that there is $j - 1$ players, and they will randomize their numbers again.

In every randomization the players randomize a permutation with probability $\frac{n!}{n^n}$, so eventually they will randomize a permutation and will reach a pure Nash equilibrium.

After every randomization the strategy is finite automata. There is a finite number of options in the randomization. Therefore the strategy is a finite automata strategy. ■

Proof of Theorem 8. We construct strategies that leads to a pure Nash equilibrium in every game with different payoffs (DP). By Lemma 17 $\lambda(DP^C) = 0$, therefore these strategies lead to pure Nash equilibrium in every generic game.

The strategies of the players can be with infinite memory, so let assume that the strategies can remember any real payoff, and that the strategies count the number of periods t.

There is a **state** of a strategy of a player:

State $k.1$: The player knows that there are at least k players, but he did not find a pure Nash equilibrium with k players.

State $k.2$: The player knows that there are at least k players, he found a pure Nash equilibrium with k players, and he remembers this equilibrium in his memory.

A player starts his play at state 1.1.

At the state $k.1$, a player will try to find an equilibrium when he assumes that there are k players, by using the step "finding a pure Nash equilibrium with n players", (see the proof of Theorem 6)[5]. If at the end of the step the player finds a pure Nash equilibrium, he changes the state of the strategy to $k.2$. Otherwise, he concludes that number of players is larger than k, and he changes the strategy state to $(k+1).1$.

At the state $k.2$, a player randomizes uniformly a natural number $1 \leq c^i \leq k+1$. Now he plays the "recognition of actions profile set" (see the proof of Theorem 4). If he concludes that there are $k+1$ players. It means that actually there are $k+1$ players or more. Therefore he changes his state to $(k+1).1$. Otherwise, if he concludes that there are less then $k+1$ players he plays his action in the pure Nash equilibrium of the k players t times (t is the number of periods till now), and stays at the state $k.2$.

One should note that if all the players use this strategy, then the states of all the players remain permanently the same, because all their conclusions after each step that was described, are identical for all the players.

Let n be the actual number of players. If all the players are at state $k.1$ for every $k \leq n$, then there is a positive probability that they will randomize k different numbers, and then they all will move to one of the states $k.2$ or $(k+1).1$ (all of them to the same one). If all the players are at state $k.2$ for every $k < n$, there is a positive probability that $k+1$ players will randomize different numbers and then they all will move to the state $(k+1).1$. So, finally, the players will get to the state $n.2$. Arriving at the state $n.2$, first the players try to find out if there are $n+1$ players, and play actions which are not an equilibrium. They play these actions a bounded by a constant (independent on t) number of times. Afterwards they play t times an equilibrium, when $t \to \infty$. So the frequency of the times that they play an equilibrium converges to 1. ■

Lemma 17 $\lambda(DP^C) = 0.$

Proof. Note that $(DP_A)^C = \{u \in \mathcal{U} : \exists i \in N,\ \exists a \neq a',\ a, a' \in A$ such that $u^i(a) = u^i(a')\}$. For every player $i \in N$, and for every pair of actions $a, a' \in A$ $a \neq a'$, let $E_{i,a,a'} = \{u \in \mathcal{U} : u^i(a) = u^i(a')\}$. $E_{i,a,a'}$ is a subspace of U_A of

[5] In the proof of Theorem 7, when the strategy save some payoff it just save its $-\log \delta$ digits after the decimal point. Now the strategies are with infinite memory, so when a strategy saves a payoff, it saves the whole real number.

dimension $n|A|-1$, therefore $\lambda(E_{i,a,a'})=0$. $(DP_A)^C = \bigcup_{i\in N,\ a,a'\in A\ a\neq a'} E_{i,a,a'}$ therefore $\lambda((DP_A)^C)=0$. $DP^C = \underset{A\in\mathcal{A}}{\cup} DP_A$ so $\lambda(DP^C)=0$. ■

Proof of Theorem 11. We define a **state** of a strategy of a player:

state $k.1$: The player knows that there are at least k players, but he did not find an $\frac{\varepsilon}{2}$-Nash equilibrium with k players.

state $k.2$: The player knows that there are at least k players, and he found an $\frac{\varepsilon}{2}$-Nash equilibrium with k players and he remembers his payoff function and the equilibrium.

The player starts his play at state 1.1.

At state $k.1$, the player finds an $\frac{\varepsilon}{2}$-Nash equilibrium when he assumes that there is k players, by the strategy "finding an ε-Nash equilibrium" (see Claim 14). The player remembers the $\frac{\varepsilon}{2}$-Nash equilibrium $(x^1,x^2,...,x^i,...,x^n)$, and his payoff function $\widetilde{u}^i_k$ in the game, where he found the equilibrium. The player changes his state to $k.2$.

Let ξ be a number small enough[6] such that for every $\frac{\varepsilon}{2}$-Nash equilibrium $x=(x^i)_{i=1}^n$ the mixed actions profile $y=(y^i)_{i=1}^n$ defined by $y^i=(1-\xi)x^i+\xi(\frac{1}{m^i},\frac{1}{m^i},...,\frac{1}{m^i})$, is an ε-Nash equilibrium.

At state $k.2$, the player plays his mixed action in the $\frac{\varepsilon}{2}$-Nash equilibrium (x^i) with probability $1-\xi$, and he plays all his actions by the uniformly distribution – with probability ξ. If his payoff is not one of the payoffs in $\widetilde{u}^i_k$, he changes his state to $k+1.1$. Otherwise he stays at the state $k.2$.

The number of players is bounded by P, and the number of actions by T, so every payoff function $\widetilde{u}^i_k$ should be saved in T^P cells of payoffs. So the automaton of the player is a finite automaton.

If the playoff of player i is not in $\widetilde{u}^i_k$, that means that the action that was played is not one of the actions in the game $\widetilde{u}_k$. Therefore, the other players will get also some payoff that is not in their payoff function. Therefore if all the players play with this strategy, then the updates of the states of all the players occur simultaneously.

Let n be the actual number of players. If all the players are at state $k.1$ for every $k\leq n$, then there is a positive probability that they will randomize k different numbers, and then they all will move to state $k.2$. If all the players are at state $k.2$ for every $k<n$, there is a probability ξ^n that all the players will play all their actions with uniform distribution. Hence there is a positive probability that the players will play some action which is not

[6]For example $\xi=\frac{\varepsilon}{4M}$ (when M is the bound of the payoffs) guaranties the requirment.

an action in $\widetilde{u}_k$, and then they will move to the state $k+1.1$. So, finally, the players will get to the state $n.2$. In the state $n.2$, the players play $(y^i = (1-\xi)x^i + \xi(\frac{1}{m^i}, \frac{1}{m^i}, ..., \frac{1}{m^i}))_{i=1}^n$, which is an ε-Nash equilibrium. The payoff function that every player save in his memory, is the actual payoff function of the game, so the players will never get a payoff which is not in their payoff function, and the players will stay at the state $n.2$ all the time. ■

Claim 18 (Analog of ε-Nash equilibrium for Theorem 6) *For every $\varepsilon > 0$, if*
(condition 1) for every player $i \in N$, α^i is the index of the player,or
(condition 2) for every player $i \in N$, α^i is the total number of the players in the game,
then there exists a completely uncoupled mapping with additional information α into finite memory strategies that leads to a ε-Nash equilibrium, except a set of games with measure smaller than ε.

Proof. Let $\nu = \nu(\varepsilon)$ be a number small enough (For example one can take $\nu = \frac{\varepsilon}{2M}$), such that for every game there exists an ε-Nash equilibrium with mixed actions which are integer multiplication of ν. Such a ν exists, because every game has Nash equilibrium, and we can approximate it by integer multiplications of ν. Now we can make a discretization of $\Delta(A^i)$ for all i, taking only the actions which are integer multiplication of ν. Denote it by $\widetilde{\Delta}(A^i)$. Then $\widetilde{\Delta}(A^i)$ is a finite set. So $\widetilde{\Delta}(A) = \widetilde{\Delta}(A^1) \times \widetilde{\Delta}(A^2) \times ... \times \widetilde{\Delta}(A^n)$ is also a finite set, and we can define a lexicographic order on $\widetilde{\Delta}(A)$.

Let us prove that each one of the two conditions is sufficient:

Condition 1: The strategy of player i is the following: He uses the following steps: "recognition of action profile set" $\rightarrow$ "recognition of payoff function" (see the proof of Theorem 4). At the end of these steps the player knows the action profile set and his payoff function. Next he plays the following way:

The player looks through all the mixed actions in $\widetilde{\Delta}(A)$ by the lexicographic order. For every action $x = (x^i, x^{-i}) \in \widetilde{\Delta}(A)$, player i plays a_1^i if x^i is an ε-best reply to x^{-i}, otherwise he plays a_2^i. If his payoff is $u^i(1)$, he remembers x^i as the ε-Nash equilibrium mixed action and moves to the next step. Otherwise he goes to the next action (by the lexicographic order).

Finally the player plays playing the ε-Nash equilibrium x^i all the time.

Similar to the proof of Theorem 6, if all the players play with this strategy, then the play converges to an ε-Nash equilibrium.

Condition 2: Also similar to the proof of Theorem 6, a player randomizes uniformly a natural number $1 \leq c^i \leq n$. If he found that there are n players he finds an ε-Nash equilibrium and play it all the time. Otherwise he randomizes c^i again. Let us call this strategy "*finding of an ε-Nash equilibrium*". By the same considerations as in Theorem 6, if all the players play by this strategy, then the play converges to an ε-Nash equilibrium. ■

Proof of Theorem 13. The proof is similar to the proof of Theorem 8. Only instead of using the step "finding a pure Nash equilibrium", we use "finding an ε-Nash equilibrium" (see Claim 14). ■

6 References

References

[1] Arslan, G., Marden, J.R., Shamma J.S., Young, H.P., 2007. Payoff based dynamics for multi-player weakly acyclic games. Decision and Control, 2007 46th IEEE Conference. 3422-3427.

[2] Foster, D.P., Kakade, S.M, 2008. Deterministic calibration and Nash equilibrium. Journal of Computer System Sciences 74. 115-130.

[3] Foster, D.P., Young, H.P., 2001. On the impossibility of predicting the behavior of rational agents. Proceeding of the National Academy of Scinces Of the U.S.A. 98. 12848-12853.

[4] Foster, D.P., Young, H.P., 2003. Learning, hypothesis testing, and Nash equilibrium. Games and Economic Behavior 45, 73-96.

[5] Foster, D.P., Young, H.P., 2006. Regret testing: learning to play Nash equilibrium without knowing you have an opponent. Theoretical Economics 1, 341-367.

[6] Germano, F., Lugosi, G., 2007. Global Nash convergence of Foster and Young's regret testing. Games and Economic Behavior 60. 135–154.

[7] Hart, S., Mas-Colell, A., 2003. Uncoupled dynamics do not lead to Nash equilibrium. American Economic Review 93, 1830-1836.

[8] Hart, S., Mas-Colell, A., 2006. Stochastic uncoupled dynamics and Nash equilibrium. Games and Economic Behavior 57, 286-303.

[9] Young, H.P., 2007. The possible and the impossible in multi-agent learning. Artifical Intelligence 171. 429-433.

[10] Young, H.P., 2009. Learning by trial and error. Games and Economic Behavior 65. 626-643.