*Simple Adaptive Strategies: From Regret Matching to Uncoupled Dynamics**

Sergiu Hart[†] Andreu Mas-Colell[‡]

May 1, 2012

INTRODUCTION

The general framework of this volume is that of *game theory*, with multiple participants ("players") who interact repeatedly over time. The players may be people, corporations, nations, computers—even genes. See the *Handbook of Game Theory* (Aumann and Hart 1994, 2002, 2004; Young and Zamir 2012).

Many of the standard concepts of game theory are *static* by their very nature. This is certainly true of the central concept of strategic equilibrium, both in its classical form of *Nash equilibrium* (Nash 1950) and in its extended form of *correlated equilibrium* (Aumann 1974). Indeed, an equilibrium situation is such that, once the players happen to find themselves in it, no player has an incentive to move away from it. However, these equilibrium concepts say nothing about how such situations are reached, i.e., about their *dynamic* basis.

*World Scientific Series in Economic Theory – Volume 4, World Scientific Publishing, 2012 (forthcoming).

[†]Center for the Study of Rationality, Department of Economics, and Institute of Mathematics, The Hebrew University of Jerusalem, Feldman Building, Givat Ram, 91904 Jerusalem, Israel. *E-mail*: `hart@huji.ac.il`; *URL*: `http://www.ma.huji.ac.il/hart`.

[‡]Department of Economics and Business, Universitat Pompeu Fabra, and Barcelona Graduate School of Economics, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. *E-mail*: `andreu.mas-colell@upf.edu`.

Yet, it is of utmost importance—theoretically as well as in applications—to study dynamic processes, and relate them to appropriate static solutions. This is so for at least two reasons. On the one hand, it is of interest in its own right to analyze processes that to some degree reflect observed patterns of behavior (we could call them "natural" dynamics). On the other hand, the significance of an equilibrium solution has to depend on how easy it is to reach: the justification of a concept that turns out to be dynamically hard to reach is shaky.

In this volume we study the connections between dynamics and equilibria. Our goal is to characterize interesting classes of natural dynamics for which convergence to Nash or correlated equilibria can be guaranteed, and classes for which it cannot (i.e., where "impossibility" results hold).

Simple Adaptive Strategies

A *strategy* in a long-term interaction provides instructions on what to do after every possible history of play. It specifies what to do in the first period, how to react in the second period to what happened in the first period, and so on. Many (in fact, most) of these strategies are extremely complex and therefore hardly practical. In this volume our interest is restricted to strategies that are *simple*—and therefore easy to implement—and *adaptive*—and therefore related to real behavioral traits. Note that we are not adding an "optimality" requirement, and so we are placing ourselves in a bounded rationality framework. In this context we are interested in studying what can be reached, and what cannot be reached, by players who use simple adaptive strategies.

Regret Matching

Our research has identified a salient class of strategies with some remarkable properties, the leading element of which is the *regret-matching* strategy. On the one hand, these strategies are based on natural measures of regret; roughly speaking, the regret is the possible increase in the payoff that one would have received had one used a different action in the past. These regret measures are used to determine one's propensity to switch to a different action: the greater the regret for an action, the greater the propensity to switch to it. As such, these are very similar to strategies obtained in the behavioral literature. They certainly qualify as instances of simple and adaptive strategies. On the other hand, these strategies, despite their deceptive simplicity, lead in the long run to behavior that is similar to that obtained from fully rational considerations, namely, correlated equilibria (Aumann 1987).

Uncoupled Dynamics

One may well ask whether it is possible to use variants of regret-based strategies to reach Nash equilibria. A long history of failure in the search for dynamic procedures guaranteeing convergence to Nash equilibria should indicate that the answer to this question is bound to be negative. And indeed it is. Yet, the reason for the failure turns out to be quite instructive. It is not due to some aspect of the adaptive property, but rather to one aspect of the simplicity property: namely, that in the strategies under consideration players do not make use of the payoffs of the other players. While the actions of the other players may be observable, their objectives and reasons for playing those actions are not known (and at best can only be inferred). This simple informational restriction—which we call *uncoupledness*—is the key property that makes Nash equilibria very hard if not impossible to reach.

Dynamics and Equilibria

At this point the connections between static equilibrium concepts and dynamics can be summarized as follows: Nash equilibrium is a "dynamically hard" concept, whereas correlated equilibrium is a "dynamically easy" concept. Indeed, regret matching and the general class of regret-based strategies—which are simple, adaptive, and natural—lead to correlated equilibria, whereas any strategies that are uncoupled cannot reasonably lead to Nash equilibria in all games.

We now provide a more detailed overview of this volume.

PART I: Correlated Equilibria

A Nash equilibrium is a situation where "each player's strategy is optimal against those of the others" (Nash 1950). In other words, it is a combination of strategies of all players such that no player can gain by unilaterally changing his strategy.

A correlated equilibrium (Aumann 1974) is a Nash equilibrium where players receive payoff-irrelevant information ("signals") before playing the game. But since this information does not affect the game itself, does it matter? Yes, it does: it may well be used by the players when making their strategic choices.

When the signals are (stochastically) independent across the players, the resulting correlated equilibrium is clearly just a standard (mixed-strategy) Nash equilibrium: the signals can only be used as a private randomizing device. But when the signals are correlated, new equilibria emerge. In case the information is public (i.e., the same signal is sent to every player, and therefore it is commonly observed and commonly known), then, after each realization of the signal, the resulting play has to be a Nash equilibrium of the original game (with no signals). Thus, what we get overall is a probabilistic mixture (i.e., a convex combination) of Nash equilibria of the original game, with weights that are precisely the probabilities of the various possible signals; call this a *publicly correlated equilibrium*. For example, let NE1 and NE2 be two distinct Nash equilibria of the original game, and assume that 60% of the days are sunny and 40% are cloudy, and this is observed by all players before they play the game. Then there is a (publicly) correlated equilibrium consisting of playing NE1 on sunny days and NE2 on rainy days, which means that NE1 is played with probability 60% and NE2 with probability 40% (in general, this behavior cannot be a Nash equilibrium of the original game, since a Nash equilibrium requires the players' choices to be independent).

Even more interesting correlated equilibria can occur when the signals are neither independent nor fully correlated. New phenomena emerge then. For example, one may get correlated equilibria whose payoffs Pareto dominate the payoffs of all Nash equilibria of the original no-signals game (e.g., in the so-called Chicken game; see Section 6 in Chapter 11).

Existence of Correlated Equilibria

Consider now a finite game; that is, there are finitely many players, each having finitely many strategies. Does a correlated equilibrium necessarily exist? The answer is clearly "yes," since Nash equilibria exist (Nash 1950), and every Nash equilibrium is also a correlated equilibrium.

While this provides a correct mathematical proof for the existence of correlated equilibria, it has a drawback. The set of correlated equilibria is characterized by a set of linear inequalities (the "incentive compatibility" constraints of the players), and is therefore a convex polytope in a finite-dimensional Euclidean space. To prove nonemptiness of such a set using fixed-point theorems (the use of a fixed-point theorem, such as that of Brouwer or of Kakutani, is needed to prove the existence of Nash equilibria) looks in such a context like "overkill." One would like to prove this result more elementarily, specifically, in a framework of linear algebra and linear inequalities.

Such an elementary proof is provided by Sergiu Hart and David Schmeidler, "Existence of Correlated Equilibria," *Mathematics of Operations Research* 14 (1989), 18–25, which is **Chapter 1** of this volume.[1] It rests on an appeal to the minimax theorem (which is essentially equivalent to the duality theorem of linear programming). To wit: given a finite n-person game, one constructs an auxiliary two-person zero-sum game, with the auxiliary row player, call it ROW, choosing an n-tuple of strategies in the original game, and the auxiliary column player, call it COL, choosing one of the incentive constraints for a correlated equilibrium. It is seen that an optimal (mixed) strategy for ROW is then precisely a correlated equilibrium in the original game; its existence then follows by applying (twice) the minimax theorem of two-person zero-sum games.

While providing the "right" proof is often a worthwhile formal exercise, in this case it had the added advantage that it eventually led to the research included in this volume. Our basic, first idea was to use the above auxiliary game and apply to it the simple dynamic of "fictitious play" (Brown 1951)—which, for two-person zero-sum games, converges and leads to optimal strategies (Robinson 1951)—in order to obtain a dynamic that leads to

[1] Another elementary proof is provided by Nau and McCardle (1990).

correlated equilibria of the original n-person game (which, as seen above, are precisely the optimal strategies of ROW). However, the resulting dynamic was unwieldy, as it required all the n players to move in a coordinated way. It took many modifications and transformations until we succeeded in "uncoupling" the players (cf. Part III) and getting the simple regret-matching strategies (see Part II).

PART II: Regret Matching

We start by considering our basic simple adaptive strategy, "regret matching," introduced by Sergiu Hart and Andreu Mas-Colell, "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica* 68 (2000), 1127–1150, which is **Chapter 2** of this volume. Consider a game that is repeatedly played over time; regret matching is defined by the following rule:

> *Regret Matching*: Switch next period to a different action with a probability that is proportional to the regret for that action, where regret is defined as the increase in the payoff had such a change always been made in the past.

That is, consider a player deciding on his action (it is convenient to use the term "action" for the choice in the underlying one-shot game, and "strategy" for the choice in the repeated game) at a certain time period. Let U be the average payoff the player has obtained up to now, and let j be the action that he played in the previous period. For each alternative action k different from j, let $V(k)$ be the average payoff the player would have obtained had he played k instead of j every time in the past that he actually played j. The *regret* $R(k)$ for action k is then defined as the amount, if any, by which $V(k)$ exceeds the actual payoff U; i.e., $R(k) = V(k) - U$ if $V(k) \geq U$, and $R(k) = 0$ otherwise. Regret matching stipulates that each action k different from the previous period's action j is played with a probability that is proportional to its regret $R(k)$, and, with the remaining probability, the same action j as in the last period is played again. (Formally, the probability of playing action k equals $cR(k)$ for every k different from j, and it equals the remaining probability $1 - \sum_{k \neq j} cR(k)$ for $k = j$; here, c is an appropriate fixed constant; see footnote 5 in Chapter 2.) We could thus say that the player contemplates, first, whether to continue to play next period the same action j as in the previous period, or to switch to a different action k. If the latter, he looks at what would have happened to his average payoff had he always replaced j by k in the past (since one is looking at the long-run average payoff, it makes sense to consider replacing j by k not just in the previous period, but also in all the other periods in the past when j was played; after all, the effect of one period becomes negligible as the number of periods increases). The player compares what he got, U, to what he would have gotten, $V(k)$. If the alternative payoff is no higher, i.e., if $V(k) \leq U$, then he has no regret for k (the regret $R(k)$ for k equals 0) and he does not switch to action k. If the alternative payoff is higher, i.e., if $V(k) > U$, then the regret for k is positive ($R(k)$ equals the increase $V(k) - U$) and the player switches to action k with a probability that is proportional to this regret.

The main result of Chapter 2 (Hart and Mas-Colell 2000) is:

> **Regret-Matching Theorem**: If each player plays a regret-matching strategy, then the joint distribution of play converges to the set of correlated equilibria of the underlying game.

The term "joint distribution of play" (also known as the "empirical distribution" or "sample distribution") refers to the relative frequency with which each combination of actions of all players has been played (more on this below). The Regret-Matching Theorem says that, for almost every history of play, the joint distribution of play converges to the set of correlated equilibria of the underlying (one-shot) game. This means that, from some time on, the joint distribution is close (i.e., within ε) to a correlated equilibrium, or, equivalently, that it is a correlated approximate (i.e., an ε-) equilibrium. The convergence here is to the set of correlated equilibria, not necessarily to a point in that set. Observe that it is the empirical distributions that become essentially correlated equilibria—not the actual play. What our result implies is that the long-run statistics of play of "regret-matchers," and that of fully rational players (who play a correlated equilibrium each period; see Aumann 1987), become essentially indistinguishable.

Clearly, regret matching, as well as its generalizations below, has the character of an adaptive strategy; in fact, it embodies commonly used rules of behavior. For instance, if all the regrets are zero ("there is no regret"), then a regret-matching player will continue to play the same action of the previous period (as in the common saying, "never change a winning team"). When some regrets are positive, actions may change—with probabilities that are proportional to the regrets: the higher the payoff would have been from switching to another action in the past, the higher the tendency is to switch to that action now. Again, this seems to fit standard behavior (recall ads such as "Had you invested in A rather than B, you would have gained X more by now. So switch to A now!", the sense of urgency being related to the magnitude of X). In the learning, experimental, and behavioral literature there are various models that bear a likeness to regret matching; see Bush and Mosteller (1955), Roth and Erev (1995), Erev and Roth (1998), Camerer and Ho (1998, 1999), Selten, Abbink, and Cox (2005), and others; probably the closest are the Erev–Roth models. Regret measures also feature in the recent neuroeconomics literature on decision-making; see Camille et al. (2004) and Coricelli, Dolan, and Sirigu (2007). Also, incorporating regret measures into the utility function has been used to provide alternative theories of decision-making under uncertainty; see Bell (1982) and Loomes and Sugden (1982).

Another interesting aspect captured by regret matching has to do with the "sluggishness" of decision-making. It has been observed that people tend to have too much "inertia" in their decisions: they stick to their current action for a disproportionately long time (as in the "status quo bias"; see Samuelson and Zeckhauser 1988 and Moshinsky 2003). Regret matching has built-in inertia: the probability of not switching (i.e., of repeating the previous period's action) is always strictly positive, and in fact regret matching generates behavior where the same action is played over and over again for long time intervals.

Regret matching is not just adaptive but also very simple (we could say unsophisticated!). Players neither develop beliefs nor reply optimally. At each time period there are "propensities" of play, which are adjusted over time in a simple and natural way. The computation that determines these propensities involves at each period a straightforward additive updating of the appropriate regret.

There are other dynamics leading to correlated equilibria: "calibrated learning" (i.e., best-reply to calibrated forecasts: Foster and Vohra 1997), and "conditional smooth fictitious play eigenvector strategy" (Fudenberg and Levine 1999). However, these dynamics require the players to make at each period a complex mathematical computation (specifically: compute an eigenvector), which, in our view, makes them neither simple nor easily interpretable in adaptive behavioral terms.

Now where does the "correlation" in the Regret-Matching Theorem come from? The answer is, of course, that it arises from the commonly observed history of play. Indeed, each player's action is determined by his regrets, which are in turn determined by the history. Thus, the exogenous signals in the definition of correlated equilibria are now *endogenously* generated—by the regret-matching strategies themselves.[2]

It is a standard hypothesis that the players observe the history of play ("standard monitoring"), which determines the joint distribution of play. Therefore, having players determine their actions based on the joint distribution of play (rather than having each player consider only his own frequencies of play, i.e., the coresponding marginal distribution) does not go beyond the commonly used monitoring framework: it is information that the players possess anyway. In fact—and this is a significant behavioral observation—people react to the joint distribution, not only to the marginals: people are very much aware of coincidences, signals, communications, and so on (even to the point of overdoing it and interpreting random phenomena and spurious correlations as meaningful). It should be emphasized that while at each stage the players randomize independently of one another, this does not imply that the joint distribution of play should be independent across players (i.e., the product of its marginal distributions), nor should it become so in the long run.

[2]While the history of play is commonly observed, it does *not* follow that one obtains *publicly* correlated equilibria; the reason is that each player plays according to his own regrets. These regrets are correlated (since they are based on the common history), but in general they are far from being fully correlated.

To summarize: reasonable models of play can—and should—take into account the joint distribution of play.

Generalized Regret-Based Strategies

The regret-matching strategy appears to be very specific: the play probabilities are directly proportional to the regrets. It is natural to enquire whether this delicate construct is necessary for the result of the Regret-Matching Theorem. What would happen were the probabilities proportional to, say, the squares of the regrets? In another direction, we could also ask for the connections between regret matching and other dynamics leading to correlated equilibria, particularly variants of conditional smooth fictitious play (e.g., Fudenberg and Levine 1999a).

This leads us to consider a large class of adaptive heuristics that are based on regrets. Specifically, instead of the switching probability being proportional to the regret, i.e., equal to $cR(k)$, we now allow this switching probability to be given by a general function $f(R(k))$ of the regret $R(k)$, provided that f is sign-preserving (i.e., $f(x) > 0$ for $x > 0$ and $f(0) = 0$) and regular (which here means Lipschitz continuous). Call the resulting strategies *generalized regret-matching* strategies, or *regret-based* strategies.

The following result is based on Sections 3.2 and 5.1 of Sergiu Hart and Andreu Mas-Colell, "A General Class of Adaptive Strategies," *Journal of Economic Theory* 98 (2001), 26–54, which is **Chapter 3** of this volume, and proved as Theorem 4.1 in[3] Amotz Cahn, "General Procedures Leading to Correlated Equilibria," *International Journal of Game Theory* 33 (2004), 21–40, which is **Chapter 6** of this volume:

> **Generalized Regret-Matching Theorem**. If each player plays some generalized regret-matching strategy then the joint distribution of play converges to the set of correlated equilibria of the underlying game.

In fact, the full class of generalized regret-matching strategies (for which

[3]Based on the master's thesis of Amotz Cahn, written under the supervision of Sergiu Hart.

the above theorem holds) is even larger; see Section 3.2 in Chapter 3 and Section 4 in Chapter 4.

As a special case, consider the family of functions $f(x) = cx^r$, where $r \geq 1$ and $c > 0$ is an appropriate constant. At one extreme, when $r = 1$, we have regret matching. At the other extreme, the limit as r goes to infinity results in the probability of switching being equally divided among those actions $k \neq j$ with maximal regret (i.e., those k with $R(k) = \max_{\ell \neq j} R(\ell)$). This yields a variant of fictitious play, which, however, no longer belongs to the admissible class of generalized regret strategies (it is not continuous), and, indeed, the result of the Generalized Regret-Matching Theorem above does not hold for it (see Section 4 in Chapter 3). Of course, the result does hold for any finite $r \geq 1$, which for very large r leads to "smooth conditional fictitious play"; see Section 4.5 in Chapter 6.

The case of the unknown game ("complete uncoupledness")

Consider now the apparently hopeless situation where each player knows initially only his own set of actions, and is informed, after each period of play, of his realized payoff. He does not know what game he is playing, that is, how many players there are and what their actions and payoffs are. Moreover, he does not even know his own payoff function—but only the payoffs he did actually receive every period. This is essentially a standard stimulus-response setup, which we called "the case of the unknown game" (also known in the literature as "payoff-based" and "radically uncoupled"; the current terminology refers to this case as "complete uncoupledness"—see Babichenko 2011).

While at each period the player knows his realized average payoff U, he cannot know his alternate payoffs $V(k)$ and his regrets $R(k)$: he knows neither what the other players did, nor what his payoff would have been had he played an alternative action k instead. Yet, we can still define a *proxy regret* measure, by using the payoffs he got when he did actually play k. In Sergiu Hart and Andreu Mas-Colell, "A Reinforcement Procedure Leading to Correlated Equilibrium," in *Economic Essays*, edited by Gerard Debreu, Wilhelm Neuefeind, and Walter Trockel, Springer (2001), 181–200, which is **Chapter 4** of this volume, we show exactly how to carry this out (it involves certain adjustments and also a small degree of experimentation). The surprising result is that convergence to correlated approximate equilibria is obtained also for proxy-regret-matching strategies.

Unconditional Regrets and Hannan Consistency

The regret for action k has been defined relative to the action j of the previous period. Consider instead a rougher measure, namely, the increase in the average payoff, if any, were one to replace *all* past plays, and not just the j-plays, by k. This yields the *unconditional regret* for action k, denoted $R^{\text{U}}(k)$ (see Section 4 (c) in Chapter 2; these regrets are also known as "external regret," with "internal regrets" for the original regrets). The resulting *unconditional-regret-matching* strategy prescribes play probabilities at each period that are directly proportional to the vector of unconditional regrets; i.e., the probability of playing k is $R^{\text{U}}(k)/\sum_{\ell} R^{\text{U}}(\ell)$, the unconditional regret of k divided by the sum of the unconditional regrets for all actions ℓ (unlike regret matching, here we do not use a constant proportionality factor c, but simply normalize the vector of unconditional regrets to get a probability vector).

A strategy of a player is said to be *Hannan-consistent* if it guarantees, for any strategies of the other players, that all the unconditional regrets vanish in the limit with probability one. We have:

> **Unconditional Regret Theorem.** Unconditional regret matching is Hannan-consistent. Moreover, if all players play unconditional regret matching, then the joint distribution of play converges to the Hannan set of the stage game.

This is Theorem B in Chapter 2 (Hart and Mas-Colell 2000). It is proved via Blackwell's Approachability Theorem (Blackwell 1956a, 1956b). The *Hannan set*[4] (see Hannan 1957, Moulin and Vial 1978, Hart and Mas-Colell

[4]Also known as the set of "coarse correlated equilibria." The Hannan set refers to the one-shot game (it is a set of joint distributions), whereas Hannan-consistency refers to the repeated game (it is a long-run property of a strategy).

2003a [Chapter 5]) is defined as the set of joint distributions of play where no player can gain unilaterally by playing the same constant action at all periods, irrespective of any signal he gets. Note the contrast with correlated equilibrium, where the no-gain test would be for changes conditional on the signal. The set of correlated equilibria is contained in the Hannan set (and the two sets coincide when every player has at most two strategies); moreover, the Hannan distributions that are independent across players are precisely the Nash equilibria of the game.

Hannan-consistent strategies have been constructed, among others, by Hannan (1957), Blackwell (1956b), Foster and Vohra (1993, 1998), Fudenberg and Levine (1995), and Freund and Schapire (1999) (many of these strategies are smoothed-out variants of fictitious play, which, by itself, is not Hannan-consistent). It appears that in comparison our unconditional regret matching is simpler.

Continuous Time

The regret-based dynamics up to this point have been discrete-time dynamics: the time periods were $t = 1, 2, \ldots$. It is natural to study also continuous-time models, where the time t is a continuous variable, and the changes in the players' actions are governed by appropriate differential equations. In Sergiu Hart and Andreu Mas-Colell, "Regret-Based Continuous-Time Dynamics," *Games and Economic Behavior* 45 (2003), which is **Chapter 5** of this volume, we show that the results carry over to this framework. In fact, some of the proofs become simpler.

Summary and Applications

For an extensive survey, with precise pointers to the relevant papers, of the results on regret matching and general regret-based strategies, the reader may consult Chapter 11 in Part IV.

Interestingly, regret matching is also successfully considered nowadays in various areas of practical application. The fact that game-theoretic adaptive procedures, such as the various regret-matching strategies in this volume, are both simple and decentralized (i.e., they can be carried out "locally" without the need to communicate with a "central authority") makes them appealing and useful in many setups where one needs to make efficient use of limited resources. Some examples: "cognitive radio," which refers to wireless transceivers that change their parameters dynamically in response to their environment (Maskery, Krishnamurthy, and Zhao 2009; Wang, Wu, and Liu 2010); traffic, congestion, and Voronoi diagrams (Arslan, Marden, and Shamma 2007; Kalam, Gani, and Seneviratne 2008); sensor networks (Krishnamurthy, Maskery, and Yin 2008); neural networks (Marchiori and Warglien 2008); statistical analysis of large datasets in medical diagnosis (Gambin et al. 2009). In many of these cases regret matching (the extremely simple unconditional regret matching, or the somewhat more sophisticated conditional regret matching) turn out to yield quite efficient results. It would be interesting to understand what exactly lies behind this apparent efficiency, as it does not follow from the general results presented in this volume (while there are correlated equilibria that are "better" than, say, Nash equilibria, in general there are also "worse" correlated equilibria).

PART III: Uncoupled Dynamics

As explained so far, regret matching and its generalizations fit the equilibrium concept of correlated equilibrium. What about Nash equilibrium? Nash equilibria belong to the set of correlated equilibria but, unless this latter set happens to be a singleton, or there is a pure strict Nash equilibrium, regret matching will not generate dynamics converging to Nash equilibrium. This, we hasten to add, came as no surprise, since dynamic procedures ensuring convergence to Nash equilibria—such as fictitious play—have only been obtained for quite restricted classes of games (such as two-person zero-sum games, potential games, 2 x 2 games). Intuitively speaking, regret matching seemed much too simple and much too adaptive to hope for a convergence result to Nash equilibria for general classes of games. But can this statement be made precise? This is the question that motivated the line of work of Part III, which we summarize now.

All the dynamic processes we have considered until now share the following characteristic, which is part of the simplicity, rather than the adaptiveness, requirement: what a player does at any moment of time does not depend on the payoff or the utility functions of the other players. It is this property, which we called *uncoupledness*, that we came to view as key and focused on in our research. It is related to the notions of "privacy-preserving" in mechanism design, "decentralized" procedures in economics, and "distributed" computations in computer science.

Uncoupledness for Deterministic Dynamics

In Sergiu Hart and Andreu Mas-Colell, "Uncoupled Dynamics Do Not Lead to Nash Equilibrium," *American Economic Review* 93 (2003), 1830–1836, which is **Chapter 7** of this volume, we defined the uncoupledness notion and established its relevance to the issue of determining Nash convergent mechanisms by means of a simple model of deterministic dynamics in continuous time. Essentially, we exhibited a class of games where any such dynamics (expressed by means of differential equations) on the state space of (continuous) action combinations, which is uncoupled, will necessarily fail to converge to (the unique) Nash equilibrium for some games.

Uncoupledness for Stochastic Dynamics

The above could not be the end of the story, however. One needed to consider the standard framework of discrete time and stochastic dynamics (which is also the context in which regret matching was originally formulated). Now, it is not difficult to see that deterministic and stochastic variants of exhaustive search ("keep looking until you find a Nash equilibrium") can be embedded into an uncoupled framework. However, exhaustive search is hardly an attractive adjustment mechanism. Thus the question was which additional considerations are important and natural (and in particular rule out exhaustive search and recover the impossibility result). These turned out to be, on the one hand, "finite recall," and on the other hand, the "speed of convergence."

An additional stimulus for our research at this point was the work of Foster and Young (2003, 2006), Germano and Lugosi (2007), Kakade and Foster (2004), Young (2004, 2009). They formulated stochastic, adaptive mechanisms (that include some form of experimention) that yield trajectories that are most of the time close to Nash equilibria. It was therefore important to clarify the relationships between these two lines of research and to find the demarcation line, so to speak, between possibility and impossibility.

In Sergiu Hart and Andreu Mas-Colell, "Stochastic Uncoupled Dynamics and Nash Equilibrium," *Games and Economic Behavior* 57 (2006), 286–303, which is **Chapter 8** of this volume, we explore the implications of recall restrictions—i.e., restrictions on how many past periods of play players remember—in a stochastic dynamic context. These restrictions certainly matter. We show, for example, that for the case where a pure strategy Nash equilibrium exists, convergence cannot be assured in the case of one-period recall. Thus, not only "best-replying to the last period" cannot lead to Nash equilibrium in general, but no uncoupled dynamic where players base their decisions only on the previous period's play can do so. But, if one increases the recall to two or more periods, then convergence to pure Nash equilibria can be guaranteed (by strategies reminiscent of exhaustive search). For the general case where all the Nash equilibria may be mixed, a more refined analysis is required, yet, again, if the recall is short, convergence is not assured, while if it sufficiently long (but still finite), then it is. However, no finite length of recall turns out to be sufficient to obtain the convergence to Nash equilibria of the period-by-period behavior probabilities; this can be obtained only within the broader context of finite memory (which leads us to the next chapter).

Uncoupledness for Finite Memory Dynamics

Recall limitations are just one way to capture the idea that the past can influence the future only through a finite number of parameters. More generally, one could appeal to the notion of finite memory, or, equivalently, strategies that can be implemented by finite automata. In the present context, this is explored in[5] Yakov Babichenko, "Uncoupled Automata and Pure Nash Equilibria," *International Journal of Game Theory* 39 (2010), 483–502, which is **Chapter 9** of this volume. The finite recall results turn out to generalize nicely: for example, to reach a pure Nash equilibrium (assuming it

[5]Based on the master's thesis of Yakov Babichenko, written under the supervision of Sergiu Hart.

exists), the number of states must be strictly larger than the number of actions.

How Long to Equilibrium?

Next, consider the important issue of the time it takes to reach Nash equilibria: can one estimate, or bound, the number of periods until an (approximate) equilibrium is reached? If that number of periods turns out to be extremely large, then there may be little use for such dynamics.

This question is addressed in Sergiu Hart and Yishay Mansour, "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures," *Games and Economic Behavior* 69 (2010), 107–126, which is **Chapter 10** of this volume.

The way one proceeds is by using a tool from theoretical computer science: communication complexity. It turns out that an uncoupled dynamic reaching an equilibrium is nothing but a so-called "distributed computational procedure." Informally, a distributed computational procedure consists of a number of agents, each one initially possessing some private information (the "inputs"), which through communication reach a situation where they all agree on a certain result (the "output"). The communication complexity is the minimal number of communication rounds that is needed to go from the private inputs to the common output.

An uncoupled dynamic reaching an (approximate) equilibrium does indeed fit this framework. The private inputs are the payoff functions (each player knows only his own), the communication phase consists of playing the game repeatedly, and the end result is the (approximate) equilibrium reached. The communication complexity of the uncoupled dynamics gives the minimum number of periods needed to reach the (approximate) equilibrium (this connection was made by Conitzer and Sandholm 2004 for two-person games; for communication complexity in general, see Yao 1979 and Kushilevitz and Nisan 1997).

In Chapter 10 it is shown that the number of periods needed to reach a Nash equilibrium—pure or mixed—can be exponential in the number of players n; when n is large, this becomes quickly unreasonably long. Since this is the time exhaustive search takes, the conclusion is that there are games where any uncoupled dynamic will take essentially as long as exhaustive search does to reach Nash equilibria. We emphasize that these exponential lower bounds apply to *any* uncoupled dynamic; additional requirements on the dynamics, such as finite recall and memory, or various incentives and rationality desiderata, can only increase the number of periods required.

At this point one may wonder whether we are perhaps asking for too much. After all, the description of the game is exponential in the number of players n; even a single player's payoff function is so (since the number of action combinations, for each one of which we need to specify a payoff, is exponential in n). It may thus seem unreasonable to expect a dynamic process to lead to equilibrium in a number of periods that is significantly shorter than the time it takes to describe the game. Yet, this is what happens for correlated equilibria: regret-based strategies do reach correlated (approximate) equilibria in a time that is polynomial in the number of players n (and one can also reach exact correlated equilibria in polynomial time; see Theorems 17 and 30 in Chapter 10). Perhaps surprisingly, this implies, in particular, that each player ends up looking only at a very small proportion of the entries in his payoff matrix. The technically savvy will note that, since correlated (approximate) equilibria are given by a linear (in n) number of inequalities, there always exist solutions whose support is polynomial in n; particular instances are the correlated approximate equilibria obtained in polynomial time by the regret-based procedures.

Indeed, impossibility results of the kind discussed here in Part III show why various past attempts to prove that certain dynamics always converge to equilibria were doomed to fail.

PART IV: Dynamics and Equilibria

Part IV ties together the various results of the previous chapters.

Adaptive Heuristics

The paper of Sergiu Hart, "Adaptive Heuristics," *Econometrica* 73 (2005), 1401–1430, which is **Chapter 11** in this volume, starts with a rough but hopefully useful classification of various dynamics into three classes: *learning dynamics, adaptive heuristics,* and *evolutionary dynamics*. Learning dynamics refers to situations where players start with certain Bayesian beliefs (on the game being played and on their opponents' strategies), which they update as the play progresses (e.g., see Kalai and Lehrer 1993 and the ensuing literature); evolutionary dynamics refers to situations where forces of selection and mutation change the populations' composition; and adaptive heuristics refers to simple rules of behavior that make a player move, roughly speaking, in seemingly payoff-improving directions. One can understand the distinctions in terms of the degree of rationality and cognitive optimization of the participants: high for learning dynamics, low for evolutionary dynamics, and in between for adaptive heuristics (see Section 2 in Chapter 11 for further details and references).

Regret matching and the generalized regret dynamics of Part II are simple rules of behavior, based on natural regret measures, and as such clearly qualify as adaptive heuristics. On the one hand, they serve by their very nature as a sort of bridge between rational and behavioral viewpoints. On the other hand, they establish a solid connection to correlated equilibria (rather than Nash equilibria). Thus simple adaptive behavior can lead in the long run to outcomes that embody full rationality (i.e, correlated equilibria; recall Aumann 1987).

Dynamics and Equilibria

More than sixty years after John Nash's Ph.D. thesis (1950), where the notion of strategic equilibrium—now known as "Nash equilibrium"—was introduced, we have learned that *there are no general, natural dynamics leading to Nash equilibrium.* This statement is proposed in the commentary of Sergiu Hart, "Nash Equilibrium and Dynamics," *Games and Economic Behavior* 71 (2011), 6–8, which is **Chapter 12** of this volume. "General" refers to dynamics that operate in all games, rather than only in some specific class of games (such as two-person zero-sum games, or two-person potential games, where such dynamics do exist). "Leading to Nash equilibrium" means that at some time the dynamic reaches a Nash equilibrium (or a neighborhood of a Nash equilibrium) and stays there from then on; therefore we do not include here the dynamics that spend most of the time (formally: $1-\varepsilon$ of the time, for small $\varepsilon > 0$) near Nash equilibria, but never remain there. Finally, we take "natural" to mean simple and adaptive.

The papers in this volume, particularly in Part III, show that the lack of dynamics to which we pointed above is not a deficiency of the existing literature, but rather a result of the inherent difficulty and even impossibility of reaching Nash equilibria. Uncoupledness—which is nothing but a simplicity property concerning the amount of information players possess—severely restricts the possibility of dynamics to lead to Nash equilibria. In Chapter 8 it is shown that, with limited (small) recall, it is impossible for such dynamics to always reach Nash equilibria, and in Chapter 10 we see that uncoupledness by itself, without any further assumptions of simplicity or otherwise, already makes, in some cases, the number of periods needed to reach Nash equilibria unreasonably large.

In short, the evidence points to Nash equilibria as being a "dynamically hard" concept, whereas correlated equilibria are, in contrast, "dynamically easy."[6]

Directions of Research

The general program to which the research presented in this volume belongs may be viewed as a two-pronged approach. On the one hand, one tries to demarcate the border between those classes of dynamics where convergence to a certain equilibrium concept can be obtained and those where it cannot (i.e., where an "impossibility" result holds). On the other hand, one looks for natural and interesting dynamics—dynamics that are related to actual behavior and yield useful insights. While significant advances have been made in both approaches (such as the research presented in this volume, and the line of work initiated by Foster and Young and discussed in connection to Chapter 8 above), and these have increased our understanding

[6]This is certainly related to—though probably not fully explained by—Nash equilibria being fixed points of nonlinear maps, whereas correlated equilibria are solutions of linear inequalities.

of the connections between dynamics and equilibria, the general picture is still far from complete.

In particular, one needs to study the convergence and trajectories properties of the various dynamics, investigate various classes of dynamics, sharpen further the distinctions between the equilibrium concepts on dynamical grounds, and use all these insights in interesting applications.

We hope bringing these papers together into one volume will facilitate further pursuit of this fascinating research.

Jerusalem and Barcelona, 2012

References

Arslan, G., J. R. Marden, and J. S. Shamma (2007), "Autonomous Vehicle-Target Assignment: A Game Theoretical Formulation," *ASME Journal of Dynamic Systems, Measurement and Control* (special issue on "Analysis and Control of Multi-Agent Dynamic Systems"), 129, 84–596.

Aumann, R. J. (1974), "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics*, 1, 67–96.

Aumann, R. J. (1987), "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

Aumann, R. J. and S. Hart (eds.) (1994, 2002, 2004), *Handbook of Game Theory, with Economic Applications*, Vols. I–III. Amsterdam: Elsevier.

Babichenko, Y. (2010), "Uncoupled Automata and Pure Nash Equilibria," *International Journal of Game Theory*, 39, 483–502. [Chapter 9]

Bell, D. E. (1982), "Regret in Decision Making under Uncertainty," *Operations Research*, 30, 961–981.

Blackwell, D. (1956a), "An Analog of the Minmax Theorem for Vector Payoffs," *Pacific Journal of Mathematics*, 6, 1–8.

Blackwell, D. (1956b), "Controlled Random Walks," in *Proceedings of the International Congress of Mathematicians* 1954, Vol. III. Amsterdam: North-Holland, pp. 335–338.

Brown, G. W. (1951), "Iterative Solutions of Games by Fictitious Play," in *Activity Analysis of Production and Allocation*, ed. by T. C. Koopmans. New York: Wiley, pp. 374–376.

Bush, R. R. and F. Mosteller (1955), *Stochastic Models for Learning*. New York: Wiley.

Cahn, A. (2004), "General Procedures Leading to Correlated Equilibria," *International Journal of Game Theory*, 33, 21–40. [Chapter 6]

Camerer, C. and T.-H. Ho (1998), "Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation," *Journal of Mathematical Psychology*, 42, 305–326.

Camerer, C. and T.-H. Ho (1999), "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, 67, 827–874.

Camille, N., G. Coricelli, J. Sallet, P. Pradat-Diehl, J.-R. Duhamel, and A. Sirigu (2004), "The Involvement of the Orbitofrontal Cortex in the Experience of Regret," *Science*, 304, 1167–1170.

Conitzer, V. and T. Sandholm (2004), "Communication Complexity as a Lower Bound for Learning in Games," in *Proceedings of the Twenty-First International Conference on Machine Learning*, ed. by C. E. Brody. New York: ACM, pp. 185–192.

Coricelli, G., R. J. Dolan, and A. Sirigu (2007), "Brain, Emotion and Decision Making: The Paradigmatic Example of Regret," *Trends in Cognitive Sciences*, 11, 258–265.

Erev, I. and A. E. Roth (1998), "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," *American Economic Review*, 88, 848–881.

Foster, D. P. and R. V. Vohra (1993), "A Randomization Rule for Selecting Forecasts," *Operations Research*, 41, 704–709.

Foster, D. P. and R. V. Vohra (1997), "Calibrated Learning and Correlated Equilibrium," *Games and Economic Behavior*, 21, 40–55.

Foster, D. P. and R. V. Vohra (1998), "Asymptotic Calibration," *Biometrika*, 85, 379–390.

Foster, D. P. and H. P. Young (2003), "Learning, Hypothesis Testing, and Nash Equilibrium," *Games and Economic Behavior*, 45, 73–96.

Foster, D. P. and H. P. Young (2006), "Regret Testing: Learning to Play Nash Equilibrium without Knowing You Have an Opponent," *Theoretical Economics*, 1, 341–367.

Freund, Y. and R. E. Shapire (1999), "Adaptive Game Playing Using Multiplicative Weights," *Games and Economic Behavior*, 29, 79–103.

Fudenberg, D. and D. Levine (1995), "Consistency and Cautious Fictitious Play," *Journal of Economic Dynamics and Control*, 19, 1065–1090.

Fudenberg, D. and D. Levine (1999), "Conditional Universal Consistency," *Games and Economic Behavior*, 29, 104–130.

Gambin, A., E. Szczurek, J. Dutkowski, M. Bakun, and M. Dadlez (2009), "Classification of Peptide Mass Fingerprint Data by Novel No-Regret Boosting Method," *Computers in Biology and Medicine*, 39, 460–473.

Germano, F. and G. Lugosi (2007), "Global Convergence of Foster and Young's Regret Testing," *Games and Economic Behavior*, 60, 135–154.

Hannan, J. (1957), "Approximation to Bayes Risk in Repeated Play," in *Contributions to the Theory of Games*, Vol. III , Annals of Mathematics Studies 39, ed. by M. Dresher, A. W. Tucker, and P. Wolfe. Princeton: Princeton University Press, pp. 97–139.

Hart, S. (2005), "Adaptive Heuristics," *Econometrica*, 73, 1401–1430. [Chapter 11]

Hart, S. (2011), "Nash Equilibrium and Dynamics," *Games and Economic Behavior*, 71, 6–8. [Chapter 12]

Hart, S. and Y. Mansour (2010), "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures," *Games and Economic Behavior*, 69, 107–126. [Chapter 10]

Hart, S., and A. Mas-Colell (2000), "A Simple Adaptive Procedure Leading to Correlated Equilibrium," *Econometrica*, 68, 1127–1150. [Chapter 2]

Hart, S., and A. Mas-Colell (2001a), "A General Class of Adaptive Strategies," *Journal of Economic Theory*, 98, 26–54. [Chapter 3]

Hart, S., and A. Mas-Colell (2001b), "A Reinforcement Procedure Leading to Correlated Equilibrium," in *Economic Essays: A Festschrift for Werner Hildenbrand*, ed. by G. Debreu, W. Neuefeind, and W. Trockel. Berlin: Springer, pp. 181–200. [Chapter 4]

Hart, S., and A. Mas-Colell (2003a), “Regret-Based Continuous-Time Dynamics,” *Games and Economic Behavior*, 45, 375–394. [Chapter 5]

Hart, S., and A. Mas-Colell (2003b), “Uncoupled Dynamics Do Not Lead to Nash Equilibrium,” *American Economic Review*, 93, 1830–1836. [Chapter 7]

Hart, S., and A. Mas-Colell (2006), “Stochastic Uncoupled Dynamics and Nash Equilibria,” *Games and Economic Behavior*, 57, 286–303. [Chapter 8]

Hart, S. and D. Schmeidler (1989), “Existence of Correlated Equilibria,” *Mathematics of Operations Research*, 14, 18–25. [Chapter 1]

Kakade, S. and D. P. Foster (2004), “Deterministic Calibration and Nash Equilibrium,” in *Learning Theory*, ed. by J. Shawe-Taylor and Y. Singer. Berlin: Springer, pp. 33–48.

Kalai, E. and E. Lehrer (1993), “Rational Learning Leads to Nash Equilibrium,” *Econometrica*, 61, 1019–1045.

Kalam, S., M. Gani, and L. Seneviratne (2008), “Fully Non-Cooperative Optimal Placement of Mobile Vehicles,” in *17th IEEE International Conference on Control Applications*, September 3–5, 2008, San Antonio, Texas, USA, pp. 1025–1030.

Krishnamurthy, V., M. Maskery, and G. Yin (2008), “Decentralized Adaptive Filtering Algorithms for Sensor Activation in an Unattended Ground Sensor Network,” *IEEE Transactions on Signal Processing*, 56, 6086–6101.

Kushilevitz, E. and N. Nisan (1997), *Communication Complexity*. Cambridge: Cambridge University Press.

Loomes, G. and R. Sugden (1982), “Regret Theory: An Alternative Theory of Rational Choice under Uncertainty,” *Economic Journal*, 92, 805–824.

Marchiori, D. and M. Warglien (2008), “Predicting Human Interactive Learning by Regret-Driven Neural Networks,” *Science*, 319, 1111–1113.

Maskery, M., V. Krishnamurthy, and Q. Zhao (2009), “Decentralized Dynamic Spectrum Access for Cognitive Radios: Cooperative Design of a Non-Cooperative Game,” *IEEE Transactions on Communications*, 57, 459–469.

Moshinsky, A. (2003), "The Status-Quo Bias in Policy Judgements," Ph.D. Thesis, The Hebrew University of Jerusalem (mimeo).

Moulin, H. and J. P. Vial (1978), "Strategically Zero-Sum Games: The Class of Games Whose Completely Mixed Equilibria Cannot Be Improved upon," *International Journal of Game Theory*, 7, 201–221.

Nash, J. F. (1950), "Equilibrium Points in n-Person Games," *Proceedings of the National Academy of Sciences U.S.A.*, 36, 48–49.

Nau, R. and K. F. McCardle (1990), "Coherent Behavior in Noncooperative Games," *Journal of Economic Theory*, 50, 424–444.

Robinson, J. (1951), "An Iterative Method of Solving a Game," *Annals of Mathematics*, 54, 296–301.

Roth, A. E. and I. Erev (1995), "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior*, 8, 164–212.

Samuelson, W. and R. Zeckhauser (1988), "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1, 7–59.

Selten, R., K. Abbink, and R. Cox (2005), "Learning Direction Theory and the Winner's Curse," *Experimental Economics*, 8, 5–20.

Wang, B., Y. Wu, and K. J. Ray Liu (2010), "Game Theory for Cognitive Radio Networks: An Overview," *Computer Networks*, 54, 2537–2561.

Yao, A. C.-C. (1979), "Some Complexity Questions Related to Distributive Computing," in *Proceedings of the Eleventh ACM Symposium on Theory of Computing*. New York: ACM, pp. 209–213.

Young, H. P. (2004), *Strategic Learning and Its Limits*. Oxford: Oxford University Press.

Young, H. P. (2009), "Learning by Trial and Error," *Games and Economic Behavior*, 65, 626–643.

Young, H. P. and S. Zamir (eds.) (2012), *Handbook of Game Theory*, Vol. IV. Amsterdam: Elsevier.