Asynchronous Choice and Markov Equilibria*

V. Bhaskar
Department of Economics
University of Essex
Colchester CO4 3SQ, UK

Fernando Vega-Redondo
Facultad de Económicas and
Instituto Valenciano de Investigaciones Económicas
Universidad de Alicante

March 1998
Current version: March 2000

Abstract

We provide a theoretical foundation for the use of Markov strategies in repeated games with asynchronous moves. If admissible strategies must display finite (arbitrarily long) memory and each player incurs a "complexity cost" which depends on the memory length required by her strategy, then every *Nash equilibrium* must be in *Markovian* strategies. If, in addition, admissible strategies have uniformly bounded memory, every *rationalizable* strategy must be *Markovian*. These results are robust to considerations of perfection and also yield interesting implications for equilibrium selection in simple contexts.

1 Introduction

Markov equilibria occupy an important position in the theory of dynamic and stochastic games. A large literature on a variety of topics ranging from industrial organization (e.g. Beggs and Klemperer (1992)) to renewable resources (e.g. Levhari and Mirman (1980)) restricts attention to Markov

*We would like to thank Sjaak Hurkens, two anonymous referees and an associate editor for helpful comments and suggestions. The second author acknowledges support by the Spanish Ministry of Education, CICYT Project no. 970131.

equilibria — players are restricted to strategies which condition only on the payoff relevant "state" variables, and are not allowed to condition upon payoff irrelevant events. There has however been little theoretical justification for such a restriction.[1] Furthermore, in the case of standard repeated games, where the "state" is constant over time (the simultaneous-action stage game is fixed), Markov equilibria are trivial and uninteresting, since they consist of repeated play of a Nash equilibrium of the stage game. This contrasts sharply with the "folk theorems" for these games (e.g. Fudenberg and Maskin (1986)) which show that a wide range of alternative behavior (and payoffs) may arise in repeated game equilibria, under relatively mild conditions.

The first interesting example of Markov equilibria in a repeated game context is due to Maskin & Tirole (1988*a*, 1988*b*), who studied a variety of models of oligopolistic competition. They showed that if firms take their decisions in an *asynchronous* manner and alternate in changing actions, the Markov equilibria of the resulting game are non-trivial. For example, in the price-setting context, phenomena such as the kinked demand curve and the Edgeworth cycle may arise in Markov equilibrium, leading to firms' profits in equilibrium bounded well above Bertrand profits. This contrast between Markov equilibrium outcomes of synchronous and asynchronous games is striking. With asynchronous moves, a richer pattern of equilibrium (Markov) behavior becomes possible since a player may react to her opponents' currently fixed actions, that are payoff-relevant. Further developments of this approach include Eaton and Engers (1990), De Fraja (1993), Jéhiel (1995) and Wallner (1999).

The present paper provides a theoretical foundation for Markov equilibria of repeated games with asynchronous moves that is based on memory costs. We consider two-player repeated games with discounting where players move in alternate periods. Players may condition their actions upon payoff irrelevant past events, but such conditioning is costly and their memory is finite (although arbitrarily large). Specifically, players' preferences over alternative configurations respond both to payoffs and memory requirements in the natural way: they are increasing in payoffs for identical memory requirements, but decreasing in these requirements for equal induced payoffs. This allows for either a scenario where complexity costs and payoffs are of comparable magnitudes or the commonly considered context where com-

[1] A recent attempt at such a justification may be found in Maskin and Tirole (1997). Other contributions in a related vein are Piccione and Rubinstein (1993) and Chatterjee and Sabourian (1999). All of them are briefly discussed in the concluding Section 5.

plexity costs are lexicographically less important than stage-game payoffs.

The setup described defines a repeated game where, given a strategy profile, a player's overall payoff depends upon two factors: the discounted sum of stage-game payoffs and memory costs. In this context, our main result is that every equilibrium of the repeated game with memory costs must involve only Markovian strategies. In fact, by making the further assumption that the memory length of strategies is uniformly bounded (at an arbitrarily high level), we are able to obtain a substantially stronger conclusion, i.e. every rationalizable strategy (in the sense of Bernheim (1994) and Pearce (1984)) must be Markovian as well.

It is worth noting at this point that we obtain similar results regardless of whether or not, for any given strategy profile, we require optimality off the induced path: in either case, the resulting strategies are Markovian. Hence our approach yields robust results. This may be contrasted with the pioneering work of Abreu and Rubinstein (1988), who introduce complexity considerations in a repeated game context with *simultaneous* moves. They focus on the Nash equilibrium concept, a given strategy taken to be preferred to an alternative more complex one if both strategies yield the same payoff against the opponent's strategy on the *equilibrium* path. They showed that this is enough to reduce quite substantially the wide range of Nash equilibrium payoffs typically resulting from standard "folk theorems". However, Kalai and Neme (1992) subsequently showed that just "a little perfection" (i.e. the requirement of equilibrium behavior at histories reached after only one deviation) is sufficient to restore the usual folk-theorem conclusions in this context.

The conclusions described are warranted only in repeated games where players move asynchronously. With synchronous moves, it is easy to see that memory considerations are not sufficient to rule out non-Markov equilibria. In this sense, our analysis may be interpreted as providing theoretical justification for Markov equilibria precisely in those repeated games where Markov equilibria may produce new and interesting behavior – recall that in repeated *simultaneous* games, any Markov equilibrium merely involves the concatenation of Nash equilibria of the stage game. For example, when the stage game is a 2×2 coordination game, it can be shown (cf. Bhaskar and Vega-Redondo (1998)) that the efficient action profile is always played along any Markov equilibrium of the corresponding asynchronously repeated game provided that (a) players are sufficiently patient; (b) the respective efficient action is less risky for at least one player. Thus, in light of Dutta's (1995) folk theorem for stochastic games,[2] the Markovian condition imposed by memory costs is seen to have striking consequences in certain simple but paradigmatic contexts.

The rest of the paper is organized as follows. Section 2 sets out our basic model of asynchronous repeated games. Section 3 presents the results. Section 4 discusses the main features of our approach as well as some of its implications. Section 5 concludes with a review of related literature and a summary.

2 The model

For expositional simplicity, we restrict our formal discussion to the case of two player interaction – cf. Subsection 4.2 below. Time (i.e. the stage of play) is indexed discretely, $t = 0, 1, 2, ...$ At every t, each player takes an action s_i^t from a finite set S_i. Short-run (or instantaneous) payoffs are given by stationary functions $\pi_i : S_1 \times S_2 \to \mathbb{R}$, indicating the payoff $\pi_i^t \equiv \pi_i(s_1^t, s_2^t)$ attained by each player i at t.

At $t = 0$, players choose their actions s_i^0 $(i = 1, 2)$ simultaneously. Thereafter, they revise their actions in alternation: player 1 at odd periods $t = 1, 3, 5, ...$ and player 2 at even periods $t = 2, 4, 6, ...$ Once the game reaches stage t, its prior history is given by the specification of past play $[(s_1^0, s_2^0), (s_1^1, s_2^1), ..., (s_1^{t-1}, s_2^{t-1})]$. This list of play contains redundancies derived from the fact that each player can revise her action only every two periods (consecutive strategy profiles can only differ in one of its two components). Therefore, it will be convenient to rely on a more compact description of histories based on the single specification, for each $t \geq 1$, of the action adopted by the player who is unable to revise it then. As explained above, this action defines the state of the system, since it is the only payoff-relevant information derived from prior history. Formally, it may be defined as follows:

$$\begin{aligned} \omega^t &= s_1^t \quad \text{if } t \text{ is even} \\ &= s_2^t \quad \text{if } t \text{ is odd.} \end{aligned}$$

If, for notational simplicity, we make $\omega^0 = s_1^0$, any t-long history $(t \geq 1)$ may be fully specified through the list $h^t = (\omega^\tau)_{\tau=0}^t$. The set of all such histories

[2] Dutta (1995) establishes that, in any (generic) stochastic game where the set of attainable payoffs displays a full-dimensionality condition and any state can be reached from any other state, every individually-rational feasible payoff is a subgame perfect equilibrium payoff provided that players are sufficiently patient. Dutta's theorem applies in particular to any infinitely repeated game with asynchronous moves and generic stage game payoffs.

may be partitioned into two subsets, $\hat{H}_1$ and $\hat{H}_2$, corresponding to the set of histories h^t ($t \geq 1$), where either it is the turn of player 1 (when t is odd) or 2 (when t is even) to revise her action. Denoting by h^0 the "empty history" prevailing at $t = 0$, the full set of histories after which either player 1 or 2 moves is respectively given by $H_1 = \hat{H}_1 \cup \{h^0\}$ and $H_2 = \hat{H}_2 \cup \{h^0\}$.

In the context described, a general strategy for player i is given by a mapping $f_i : H_i \to \Delta(S_i)$ specifying the probability $f_i(h^t)(s)$ with which player i chooses each action $s_i \in S_i$ after every possible history h^t such that it is her turn to move. Let F_i denote the set of strategies of player i.

As explained, the essential postulate underlying our approach is that memory is costly and thus players will aim at economizing on it (see Subsection 4.1 for a motivation). In order to evaluate the memory requirements of different strategies, we proceed as follows. First, we define first the zero-memory strategies:

$$F_i^0 \equiv \left\{ \begin{array}{c} f_i \in F_i : \left[h^t = (\omega^\tau)_{\tau=0}^t,\ \tilde{h}^{t'} = (\tilde{\omega}^\tau)_{\tau=0}^{t'} \in H_i,\ \omega^t = \tilde{\omega}^{t'} \right] \\ \Rightarrow f_i(h^t) = f_i(\tilde{h}^{t'}) \end{array} \right\}.$$

Verbally, F_i^0 stands for the strategies of player i which are only responsive to the current state of the system, i.e. the fixed action to which her opponent is committed from the previous period or empty (initial) history. As customary, they are labelled *Markov strategies*.

As player i considers tailoring her action to (payoff-irrelevant) history of progressively more protracted length, the following strategy subsets obtain for each $k = 1, 2, ...$:

$$F_i^k \equiv \left\{ f_i \in F_i : \left[\begin{array}{c} h^t, \tilde{h}^{t'} \in H_i,\ \omega^{t-s} = \tilde{\omega}^{t'-s}, \\ s = 0, 1, ...k;\ t, t' \geq k \end{array} \right] \Rightarrow f_i(h^t) = f_i(\tilde{h}^{t'}) \right\}. \quad (1)$$

For each $k = 1, 2, ...$, the strategy set F_i^k allows player i's action to depend on the partial history of length k. If a strategy belongs to F_i^k, it prescribes the same action at two histories h^t, $\tilde{h}^{t'}$ such that the last k observations are identical at both these histories. Note that this is also the case when $t \neq t'$, i.e. the two histories concerned are of different length. This reflects the assumption that a player also needs memory to keep track of calendar time.

If a strategy f_i needs exactly k-period memory to be carried out (i.e. $f_i \in F_i^k \backslash F_i^{k-1}$), we denote $m(f_i) \equiv k$. We shall assume that players must rely on strategies displaying finite (although arbitrarily long) memory. That is, we postulate:

(B) The set of admissible strategies for player $i = 1, 2$ is given by $\tilde{F}_i \equiv \bigcup_{k=0}^{\infty} F_i^k$.

The above condition posits no finite upper bound on the memory requirements displayed by admissible strategies. This is enough for our main result concerning Nash equilibria of the repeated game. For the stronger conclusion pertaining to rationalizable strategies, we need to strengthen Condition (B) as follows:

(B)$'$ There exist some *finite* ν_i $(i = 1, 2)$ such that the set of admissible strategies for player i is given by $\tilde{F}_i \equiv \bigcup_{k=0}^{\nu_i} F_i^k$.

Given any strategy pair $f = (f_1, f_2)$, we can uniquely define a probability measure over resulting action paths. Denote by $\sigma^t(f) \in \Delta(S_1) \times \Delta(S_2)$ the probability vector over players' actions at t induced by strategy profile f. This allows us to define the expected discounted flow of stage payoffs for any player i as follows:

$$V_i(f) = \sum_{t=0}^{\infty} \delta^t \, \mathbf{E}\left[\pi_i(\sigma^t(f))\right],$$

where $\mathbf{E}[\cdot]$ stands for the expectation operator.

We are now ready to specify players' preferences over strategy profiles $f = (f_1, f_2) \in \tilde{F} \equiv \tilde{F}_1 \times \tilde{F}_2$. There are three dimensions which we take into account. The first is the discounted sum of stage game payoffs which accrue to the player along the path induced by any such profile f. The second is the memory requirement of player i's strategy, $m(f_i)$. The final dimension is the expected discounted sum of stage game payoffs induced by f after arbitrary histories, i.e. histories that may be reached only if one or both players deviate from f. We shall define two different classes of preference relations, labelled Ξ^N and Ξ^P, reflecting these dimensions in different ways.

The class Ξ^N incorporates only the first two dimensions, and does not take into account payoffs after histories which are off the path induced by f. This approach is hence similar to that adopted by Abreu and Rubinstein (1988).

The second class of preference relations, Ξ^P, also takes into account repeated game payoffs which arise after deviations from f. In particular, preferences belonging to this class do *not* require that memory considerations are infinitely more important than considerations of subgame perfection. This is akin to the approach of Kalai and Neme (1992).

Denote by $\succcurlyeq_i$ the binary preference relation of each player $i = 1, 2$. It will be assumed to belong either to the class Ξ^N or to the class Ξ^P. Formally, if it belongs to Ξ^N, it must verify the following axiom:

(N) Suppose that $f,\ f' \in \tilde{F}$ satisfy *both*:
(i) $V_i(f) \geq V_i(f')$;
(ii) $m(f_i) \leq m(f'_i)$.
Then, $f \succcurlyeq_i f'$. If, furthermore, (i) or/and (ii) apply strictly, $f \succ_i f'$. [3]

Axiom (N) simply requires that if a strategy profile f is no worse than another profile f' pertaining to *both* discounted payoffs and memory costs, then f is weakly preferred to f'; furthermore, if f is also strictly better than f' in one of those two respects, it is strictly preferred as well.

Note that Axiom (N) is quite weak and does not therefore specify the preference relation completely, i.e. it is consistent with a wide range of possible players' preferences. It allows, in particular, for the possibility that discounted payoffs and memory costs are of comparable (i.e. commensurable) importance. But, in addition, it admits as well the possibility that either of those two dimensions is "infinitely more important" than the other. Thus, for example, it is compatible with those commonly postulated *lexicographic preferences* where memory costs are of secondary importance – that is, they are relevant only when discounted payoff considerations are identical between two alternative outcomes.

The second class of preferences to be considered, Ξ^P, allow us to incorporate considerations of perfection, i.e. the requirement of optimality after *every* possible history. Abusing previous notation, denote by $V_i(f, h^t)$ the continuation payoff earned by player i after history h^t when the strategy profile is f. Preferences in Ξ^P are characterized by the following alternative axiom:

(P) Suppose that $f,\ f' \in \tilde{F}$ satisfy *both*:
(i) $V_i(f, h^t) \geq V_i(f', h^t)$, $\forall h^t \in H$;
(ii) $m(f_i) \leq m(f'_i)$.
Then, $f \succcurlyeq_i f'$. If, furthermore, at least one of the following applies:
(a) $\exists\ h^t \in H : V_i(f, h^t) > V_i(f', h^t)$;
(b) $m(f_i) < m(f'_i)$
then, $f \succ_i f'$

The essence of axiom (P) is that a player can be ensured to prefer one strategy profile to another only if the former does better in payoff terms than the latter after all histories, including those which are not reached when the strategy profile is played. In other words, a player may prefer to use a

[3] As customary, $f \succ_i \tilde{f}$ means: $f \succcurlyeq_i \tilde{f}$ and $\neg(\tilde{f} \succcurlyeq_i f)$.

strategy with a longer memory requirement even if there is shorter-memory strategy which does as well in the absence of any deviations, provided the former does better in the event of some deviations.

Let $\Xi = \Xi^N \cup \Xi^P$. Note that we do not have to assume that the preference relation $\succcurlyeq_i$ is complete or transitive.[4] Our analysis is completely general, and applies to *any* preference relations $\succcurlyeq_i$ in the class Ξ and *any* Nash equilibrium induced by them. (In fact, if the stronger condition (B)$'$ applies, they apply to any rationalizable strategy.) However, observe that if players' preferences belong to the class Ξ^P, the induced equilibrium embodies a notion of perfection, a feature not necessarily displayed if preferences belong to the class Ξ^N. To emphasize this distinction, the equilibrium will be labelled Perfect* Equilibrium in the former case whereas we shall refer to it as a Nash* Equilibrium in the latter case. More formally, we propose:

Definition 1 *Given any pair of preferences* $(\succcurlyeq_1, \succcurlyeq_2) \in \Xi^2$, *a strategy profile* $f^* = (f_1^*, f_2^*) \in F$ *is said to be an* equilibrium *if, for each* $i = 1, 2$ *and any* $f_i \in F_i$, *we have* $f^* \succcurlyeq_i (f_i, f_j^*)$, $j \neq i$. *If each* $\succcurlyeq_i \in \Xi^N$ $(i = 1, 2)$, f^* *is called a* Nash* Equilibrium (N*E). *If each* $\succcurlyeq_i \in \Xi^P$ $(i = 1, 2)$, f^* *is called a* Perfect* Equilibrium (P*E).

It is worth noting that, in general, the notion of a P*E is neither stronger nor weaker than the notion of a N*E. It is obviously not weaker since the latter may prescribe sub-optimal behavior at unreached information sets. It is not stronger either since, in the case of P*E, memory considerations are *certain* to be decisive in discarding a particular strategy only if there is an alternative one that does as well after *all* histories.[5] However, the results are quite similar irrespective of the solution concept we employ. As explained in the introduction, this stands in contrast with the existing literature which introduces complexity costs in repeated games.

[4] We thank an anonymous referee for pointing this out.

[5] Note also that the above definition of P*E is neither stronger nor weaker than the standard concept of Subgame Perfect Equilibrium (SPE). That a SPE need not be a P*E is clear, since the former ignores memory considerations. On the other hand, a P*E may allow for the existence of unilateral gains in intertemporal discounted payoffs if this requires using a longer memory strategy. If one assumes that memory costs are lexicographically secondary in importance, any P*E must obviously be a SPE; however, we do not make this assumption since we seek to prove our main results under weak assumptions.

3 Results

Abstracting from memory costs, a Subgame Perfect Equilibrium where players only use strategies in F_i^0 is typically called a *Markov Perfect Equilibrium* (MPE). As well known (see Fudenberg and Tirole (1991, p. 504), there always exists a MPE in finite stochastic games, i.e. in games where players have finite action sets and the state space is finite. Obviously, in our context, a MPE is always an equilibrium for any preferences in Ξ. Hence we may conclude that the game with memory costs always has an equilibrium (both a N*E and a P*E). In fact, our main theorem shows that such MPE are the *unique* equilibria in our case for any given players' preferences belonging to Ξ.

Theorem 1 *Consider any alternating-move game as described above and let (B) hold, with players' preferences belonging to* Ξ^N *(resp.* Ξ^P*). If* $f^* = (f_1^*, f_2^*)$ *is a N*E (resp. P*E), then* $f_i^* \in F_i^0$ *for each* $i = 1, 2$.

Proof. Given preferences $\succsim_i \in \Xi$, let $\beta_i(f_j) \subset \tilde{F}_i$ stand for the set of player i's strategies that are a best response (in terms of $\succsim_i$) to any given opponent's strategy $f_j \in \tilde{F}_j$ $(i \neq j)$. It is enough to show that

$$[m(f_j) = k \geq 1] \Rightarrow [\forall f_i \in \beta_i(f_j),\ m(f_i) < k]\,. \tag{2}$$

Consider any $f_j \in \tilde{F}_j$ with $m(f_j) = k \geq 1$, and let $s^{t,k-1} = (s_j^{t-1}, s_i^{t-2}, ..., s_h^{t-k+1})$ where $h = i$ if k is odd, and $h = j$ if k is even. Given any such strategy for player j, let $W_i(\cdot)$ be the associated value function of player i in any period such that it is player j's turn to move. Since player j conditions only on actions of the last k periods, this value function only depends upon player $i's$ current action, s_i^t, and $s^{t,k-1}$. Hence player i's decision problem at each date t may be formulated as follows:

$$\max_{s_i \in S_i} u(s_i, s_j^{t-1}) + \delta W_i(s_i, s^{t,k-1})$$

By standard arguments from dynamic programming, there exists a solution to this maximization problem, $s_i(s^{t,k-1})$, where player i's action only depends upon the $k-1$ period history. Therefore, in response to the given f_j, there is some strategy which maximizes player i's repeated game payoffs and utilizes memory $k-1$ or less. Such a strategy is clearly strictly better than any strategy which uses more memory, by condition (N) or (P). The proof is thus complete. ■

Assume now the stronger condition (B)$'$ on the set of admissible strategies, so that we have fixed (but arbitrarily large) upper bounds ν_1 and ν_2 on memory lengths.[6] This leads to a substantially stronger conclusion.

Theorem 2 *Consider any alternating-move game as described above and let (B)$'$ hold, with players' preferences belonging to Ξ^N (resp. Ξ^P). If f_i is a rationalizable strategy for any player $i \in \{1,2\}$, then $f_i \in F_i^0$.*

Proof. First, simply adapt previous notation and define, associated to any subset $D_j \subset F_j$, the set

$$\beta_i(D_j) \equiv \{f_i \in F_i : \exists f_j \in D_j \text{ s.t. } f_i \in \beta_i(f_j)\} \qquad (i,j = 1,2;\ i \neq j),$$

and

$$m(D_j) \equiv \sup\{m(f_j) : f_j \in D_j\}.$$

A direct counterpart of the argument used in the proof of Theorem 1 implies that

$$[D_j \subset F_j,\ m(D_j) \geq 1] \Rightarrow m(\beta_i(D_j)) < m(D_j). \tag{3}$$

To complete the proof, we recall the concept of rationalizable strategies (Bernheim (1984) and Pearce (1984)). It requires constructing a sequence of subsets $\{D_i^q\}_{q=0}^{\infty}$ for each $i = 1,2$ as follows: First, $D_i^0 = F_i$, $i = 1,2$. Next, for each $q = 1,2,...$, we make $D_i^q = \beta(D_j^{q-1})$. Then, the set of rationalizable strategies for each player i is given by $R_i \equiv \bigcap_{q=1}^{\infty} D_i^q$. It is immediate to see that, in view of (3), $R_i \subseteq F_i^0$, which completes the proof. ■

The power of this result is that only rationalizability needs to be invoked in order to ensure that players play Markov strategies. In consequence, one can adapt well-known results in the learning literature to conclude that any payoff-monotonic evolutionary system (cf. Nachbar (1990)) or any adaptive learning process (Milgrom and Roberts (1991)) would lead players to rely only on Markov strategies. We refer the reader to Bhaskar and Vega-Redondo (1998) for the details.

4 Discussion

We now discuss a number of theoretical issues raised by the above results.

[6] In fact we only need assume that condition (B)$'$ applies for some player i.

4.1 Costly memory and strategic complexity

Our approach to measuring strategic complexity bears comparison with the literature on finite automata, as exemplified by Kalai and Stanford (1988), Abreu and Rubinstein (1988), Lipman and Srivastava (1990) and Kalai and Neme (1992). While complexity considerations are generally acknowledged to be important, there seems to be no universally accepted measure of strategic complexity. In particular, radically different results are obtained depending upon whether one uses a measure based on the number of states (Abreu and Rubinstein) or if the transitions between states are also taken into account (Lipman and Srivastava).

In essence, both our complexity measure as well as all of those referred above link the complexity of a strategy f_i to the size of the coarsest history partition for which f_i remains a measurable mapping. Of course, any such partition can always be used to define a similarity relationship among the set of all possible histories (i.e. two histories may be called "similar" if they belong to the same element of the partition). In most of the literature, however, such a notion of similarity only arises *indirectly* as a by-product of the strategy formulation. In contrast, our memory-based approach to measuring strategic complexity proceeds in the opposite way. It first postulates a primitive (time-based) notion of similarity across histories, only then defining the collection of possible partitions that may be used to assess strategic complexity. This precludes, for example, that two histories that differ in recent periods could be judged similar (i.e. lie in the same element of the partition) whereas two that differ only in terms of actions taken very far back might not.

To illustrate the point, let the stage game be the Prisoners' Dilemma and focus on the familiar grim-trigger strategy. Consider two equal-length histories, h and h', such that in both of them players have played C throughout except in period one; in period one, both players have chosen C under history h, whereas one player has chosen D under h'. The grim-trigger strategy outputs different actions at these two histories, irrespective of the length of these histories. Consequently, it requires *unbounded* (implicitly costless) memory, even though it is typically conceived as a simple strategy (it just involves two elements in the induced partition).

Suppose now that players are restricted to strategies with some *finite* memory k, where k could be arbitrarily large. Consider the "counterpart" of the grim trigger strategy in the finite memory context, which responds to a constant k-string $CCCC...CC$ with C, but plays D after any other k-string. In the absence of memory costs, a pair of such strategies can define a Nash equilibrium along the equilibrium path, provided that the players are sufficiently patient. However, it does *not* induce equilibria at all histories off the equilibrium path. In particular, it is not optimal to play D after the k-string $DCC....CC$. Hence finite memory, in conjunction with perfection considerations, suffices to rule out grim trigger type equilibria. Put differently, a grim-trigger strategy type equilibrium can be subgame-perfect in an asynchronous choice game only because memory is *literally* infinite. We believe that this suggests an additional reason why trigger-like strategies might be questionable in asynchronous choice games, despite their formal simplicity.[7]

In any case, we want to argue that if, say, the passage of time impairs a player's ability in discriminating among (or reacting to) past events, it is natural to introduce a time-based notion of similarity among histories and require that any given strategy, whatever its complexity, react in the same fashion to similar histories. And then, if the implementation of the strategy involves a cost that depends on how fine is the induced partition, the player should prefer (*ceteris paribus*) to let her behavior depend only on more recent history and thus economize on those costs. Somewhat related ideas can be found in the interesting paper of Dow (1991) where, in a model of search, the implications of limited memory are also formalized through similarity partitions on the set of past observations. On the other hand, in the context of repeated games, the implications of bounded memory have been studied by Sabourian (1991, 1997), who explores how the classical folk theorems fare under this restriction.

4.2 The role of asynchronicity

For expositional simplicity, our discussion has focused on a setup involving only two players who move in alternation. However, it is straightforward to verify that the above result extends to any context involving a finite number

[7]This example clarifies the *distinct* role played by the two assumptions, bounded memory (assumption (B)) and memory costs (assumption (P) or (N)), underlying our proof. Suppose that player i's strategy f_i has $m(f_i) = k \leq m(f_j)$. Since player $j(\neq i)$ will not condition his behavior on what happened $k+1$ periods ago, f_i can only be optimal (even abstracting from memory costs) if player i is indifferent between f_i and another strategy f_i' which has memory length $k-1$ or less. The finite counterpart of the grim trigger strategy does not achieve such indifference after some histories, and hence cannot be part of an (ordinary) Subgame-Perfect equilibrium. Of course, in the absence of memory costs, one can still have other non-Markov equilibria which achieve such indifference, typically via randomization. It is at this point that memory costs (i.e. (P) or (N)) ensure that f_i' is strictly preferred to f_i, thus ensuring that non-Markov strategies cannot be optimal.

of players who (almost surely) never receive a simultaneous opportunity to revise their actions. For example, one could postulate, as in Lagunoff and Matsui (1997), that revision opportunities for each player arrive through an independent Poisson process in continuous time.

Thus, the feature of the model that does play a critical role in our analysis is the assumption that players' choice is (a.s.) asynchronous. In simultaneous-move repeated games, memory considerations do *not* imply Markov strategies. Specifically, the proof of (2) above does not apply in this case since, if some player i conditions his strategy on the action profile prevailing k periods ago, it may in general be optimal for player $j \neq i$ to also do so. Hence the induction argument underlying the argument does not hold any longer.

For a specific example, consider the repeated Prisoners' Dilemma with *simultaneous* moves, where players preferences are given by an element of Ξ^P. Clearly, the only Markov Perfect Equilibrium of this repeated game is one where the players choose D every period irrespectively of history. However, if we assume that the cost of one period memory is small relative to the payoff benefit from responding optimally at histories which arise from a single deviation from the equilibrium path, one can support cooperation via a P*E. Consider the following strategy which requires one period memory:

Play C ("cooperate") at $t = 1$ or if the action profile at $t-1$ was (C, C).

Play D ("defect") otherwise.

The profile where both players choose this strategy is a P*E which supports the play of (C, C) in every period, provided that δ is sufficiently large. This example illustrates that even with memory costs, one can sustain non-Markov behavior with synchronous moves, provided that perfection considerations are incorporated into preferences.

4.3 Time dependence

In our framework, if any strategy f_i of player i displays some time dependence (i.e. the behavior prescribed at some point depends on the game being at a particular t) there is *no* finite k such that $f_i \in F_i^k$. Thus, from an strictly formal viewpoint, such a strategy should be taken to require infinite memory. However, if the range of periods where such time dependence may occur is finite, it is intuitive that the strategy in question should be conceived as reflecting finite memory.

Such a paradoxical state of affairs is allowed within our model only for the sake of notational simplicity. Our results would still go through under an alternative, more natural, formulation of the memory costs incurred by time-dependent strategies. To illustrate matters, suppose that a particular strategy f_i agrees with a Markov strategy in every period except period z, at which date it plays differently. Heuristically, we would insist that $m(f_i) = z$.

Our framework can be simply modified to formally model this idea. Let $\bar{\omega}$ denote the "null action" and given any history $h^t = (\omega^\tau)_{\tau=0}^t \in H_i$ define the augmented history $\check{h}^t = (\check{\omega}^\tau)_{\tau=-\infty}^t$ where $\check{\omega}^\tau = \omega^\tau$ for $0 \leq \tau \leq t$ and $\check{\omega}^\tau = \bar{\omega}$ for $\tau < 0$. Let $\check{H}_i$ be the set of augmented histories, and define each player's pure strategy as a map from $\check{H}_i$ to his stage game action set. Define the set of k-memory strategies, $\check{F}_i^k$, as in (1) with the difference that H_i is replaced by $\check{H}_i$, i.e. each history is replaced by the corresponding augmented history. Clearly, the time dependent strategy f_i now has finite memory z. Nevertheless, our results still go through since the best response to a time dependent strategy with k-period memory continues to be a strategy with $(k-1)$-period memory or less.

4.4 Markov strategies and equilibrium selection

To illustrate that memory conditions have sharp predictive power in asynchronous repeated games can we consider the following coordination game with stage game payoffs as given by the following table:

1 \ 2	A_2	B_2
A_1	a_1, a_2	d_1, c_2
B_1	c_1, d_2	b_1, b_2

Table 1

For each $i \in \{1, 2\}$, assume $a_i > c_i$, $b_i > d_i$, and $a_i > b_i$, so that $A = (A_1, A_2)$ and $B = (B_1, B_2)$ are strict Nash equilibria of the stage game, and A is the Pareto-efficient one. Further assume that for *some* player $i \in \{1, 2\}$, we have $a_i - c_i > b_i - d_i$ — in this case we say that A_i is a *safe action* for player i. This condition is substantially weaker than the requirement that A is risk dominant, as in Harsanyi-Selten (1988), which requires that $(a_1 - c_1)(a_2 - c_2) > (b_1 - d_1)(b_2 - d_2)$. In Bhaskar and Vega-Redondo (1998), we prove the following selection result

Proposition 1 *Consider any alternating-move game where the stage payoffs are as described in Table 1 and $a_i - c_i > b_i - d_i$ for some player i. Then, there exists some $\delta_0 < 1$ such that, if $\delta \geq \delta_0$, every Markov profile f^* which* *is rationalizable*[8] *induces a path of play* $(s_1^{t*}, s_2^{t*})_{t=0}^{\infty}$ *such that* $s_i^{t*} = A_i$ *for each* $i = 1, 2$ *and all* $t \geq 2$.

Even though the proof of Proposition 1 is somewhat involved, its main intuition is as follows. A Markov pure strategy for player i consists of a triple which specifies an initial action at $t = 0$, an action when the state is $\omega_i = A_j$, and an action at the state $\omega_i = B_j$. Consider any player i whose action A_i is safe. In this case, for *any* Markov strategy of player j, player i must play A_i at the state $\omega_i = A_j$. Indeed, this is so even if j plays the action B_j at every state. In this latter case, player i cannot avoid a transition to the inferior configuration B. However, she prefers that j initiate such a transition rather than initiating it herself. If i initiates such a transition, her payoff over the next two periods is $c_i + \delta b_i$, whereas if she lets j initiate the transition, her payoff is $a_i + \delta d_i$ (in either case, the payoffs after two periods are identical). The latter payoff is strictly greater for any δ if A_i is a safe strategy. Hence i will not initiate the transition from the efficient equilibrium to the inefficient one if she plays a strategy which is a best response to *any* Markov strategy of j. In view of this, j will also not initiate a transition, and this suffices to ensure that the action profile A is absorbing. This in turn implies that if players were playing the inefficient action profile, i will be willing to initiate a transition to the efficient one if she is sufficiently patient, secure in the knowledge that the other player will follow.[9]

[8] Here, we rely on the extensive-form version of rationalizability (see Bernheim (1984, p. 1022) or Pearce(1984, p. 1039)).

[9] After the first version of our paper was completed, we came across the work of Haller and Lagunoff (1997), who analyze Markov perfect equilibria in asynchronous choice games. They also have an efficient equilibrium selection result for 2×2 common interest games. However, their selection result is weaker than Proposition 1 in several respects. First, it relies on an equilibrium notion (MPE), whereas our result requires the weaker notion of rationalizability. Second, the payoff condition they require is also stronger — they require that $d_1 > c_1$ and $d_2 > c_2$, whereas we only demand that $a_i + d_i > b_i + c_i$ for some $i \in \{1, 2\}$.

Our selection result is also related to recent work by Lagunoff and Matsui (1997), who focus on *pure* coordination games (where players obtain *identical* payoffs at every action profile) and show that, if players are patient enough, the only equilibrium payoff of the repeated game is the efficient one. Their result depends critically on the fact that a pure coordination game has non-generic payoffs, since Dutta's (1995) folk theorem (recall Footnote 2) applies to generic setups. In contrast, our analysis (i.e. a conjunction of Theorem 2 and Proposition 1) demonstrate that memory considerations can have powerful effects in reducing the multiplicity of equilibria in *generic* scenarios, even if one only relies on such a weak solution concept such as rationalizability.

5 Summary and related literature

The main objective of this paper has been to provide a theoretical rationale for Markov equilibria in repeated games with asynchronous moves. Although the restriction to Markov equilibria has been ubiquitous in applications, there has hitherto been relatively little in terms of theoretical justification for this assumption.

An important exception is Maskin and Tirole (1997), who develop and refine the notion of Markov perfect equilibrium in a large class of dynamic games, and show that the concept is robust to small perturbations in payoffs. These authors also provide the following learning model to justify MPE. Players are drawn from a large population and randomly matched to play a dynamic game. In between rounds of play for the dynamic game, each player observes the information partition of the other players. Players then have the option to increase (but not to decrease) the complexity of their strategies, as measured by the extent to which the depend on past history. Each increase in complexity incurs a cost, which is sunk. Then, Maskin and Tirole show that if players are sufficiently patient, they will prefer to start with simple strategies. For this allows them the option to upgrade their complexity subsequently, but only if it becomes worthwhile. Thus, since all players start with Markov strategies, they continue to play Markov strategies thereafter and the option to increase strategic complexity is never used.

In comparison with Maskin and Tirole (1997), the rationale for Markov equilibrium offered in this paper is more modest since it applies only to asynchronous move games. However, for this type of games, our rationale is rather sharp and robust: if players economize on memory, *any* equilibrium, whether Nash or subgame perfect, is necessarily in Markov strategies (Theorem 1). Our formal argument applies to rationalizable strategies as well if memory requirements are uniformly bounded (cf. Theorem 2) and hence indicates that a fairly wide range of learning dynamics (where players may both upgrade or downgrade the memory requirements of their strategies) may be used to provide support for Markov strategies.

Two other papers by Piccione and Rubinstein (1993) and Chatterjee and Sabourian (1999) are also related to our present concerns, albeit more distantly. Piccione and Rubinstein study an infinitely repeated strategic context where the constituent stage game is one in extensive-form with perfect information. Their focus is on the equilibria of the repeated game when strategies are to be implemented through a finite automaton which bears a cost that depends on its underlying number of states. The authors show that any equilibrium of the corresponding "automata game" induces a path of play where some particular Nash equilibrium of the stage game is constantly being played. Thus, as in our case, complexity costs imply that agents must play Markovian strategies at equilibrium. Chatterjee and Sabourian study an intertemporal unanimity bargaining game played by many players. As well known, this game displays a large multiplicity of equilibria if more than two players are involved, and any individually rational outcome can be obtained at some equilibrium. To address the multiplicity problem, the authors introduce complexity costs that, as in here, are linked to the memory length required by the different strategies (although they are lexicographically less important than ordinary payoffs). These costs are shown to induce a stringent bound on possible delay at equilibrium and, if some infinitesimal perturbation is added to the model, even shrink down the equilibria of the model to the unique subgame-perfect one in stationary (Markov) strategies. Thus, again in this quite different context, it turns out that complexity costs select for Markovian equilibria.

References

Abreu, D. & A. Rubinstein (1988), "The structure of Nash equilibria in repeated games played by finite automata", *Econometrica* **56**, 1259-1281.

Beggs, A., and P. Klemperer (1992), "Multi-period competition with switching costs", *Econometrica* **60**, 544-565.

Bernheim, D. (1984): "Rationalizable strategic behavior", *Econometrica* **52**, 1007-28.

Bhaskar, V. and F. Vega-Redondo (1998): "Asynchronous choice and Markov equilibria: Theoretical foundations and applications", mimeo, University of Essex and Universidad de Alicante.

Chatterjee, K and H. Sabourian (1999): "Multiperson bargaining and strategic complexity", *mimeo*, Pennsylvania State University and Cambridge University.

De Fraja, G. (1993), "Staggered vs synchronized wage setting in oligopoly", *European Economic Review* **37**, 1507-1522.

Dow, J. (1991), "Search decisions with limited memory", *Review of Economic Studies* **58**, 1-14.

Dutta, P. (1995): "A folk theorem for stochastic games",*Journal of Economic Theory* **66**, 1-32.

Eaton, J. , and M. Engers, (1990), "Intertemporal price competition", *Econometrica* **58**, 637-659.

Fudenberg, D. & D. Kreps (1993): "Learning mixed equilibria", *Games and Economic Behavior* **5**, 320-67.

Fudenberg, D. & D. Levine (1998): *The Theory of Learning in Games,* Cambidge: MIT Press.

Fudenberg, D. & J. Tirole (1991): *Game Theory,* Cambridge (Mass.): MIT Press.

Haller, H., and R. Lagunoff (1997), "Markov equilibrium in asynchronous choice repeated games", mimeo, Georgetown.

Jéhiel, P. (1995), "Limited horizon forecast in repeated alternate games", *Journal of Economic Theory* **67**,497-519.

Kalai, E., and A. Neme.(1992), "The strength of a little perfection", *International Journal of Game Theory* **20**, 335-355.

Kalai, E., and W. Stanford (1988), "Finite rationality and interpersonal complexity in repeated games", *Econometrica* **56**, 397-410.

Lagunoff, R. & A. Matsui (1997): "Asynchronous choice in repeated coordination games", *Econometrica* **65**, 1467-77.

Levhari, D. and L. Mirman (1980), "The great fish war", *Bell Journal of Economics* 12, 322-344.

Lipman, B., and S. Srivastava (1990), "Informational requirements and strategic complexity in repeated games", *Games and Economic Behavior* **2**, 273-290.

Maskin, E. & J. Tirole (1988*a*): "A theory of dynamic oligopoly, I: overview and quantity competition with large fixed costs", *Econometrica* **56**, 549-70.

Maskin, E. & J. Tirole (1988*b*): "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles", *Econometrica* **56**, 571-600.

Maskin, E. & J. Tirole (1997), "Markov equilibrium I: Observable Actions", mimeo, Harvard University and Toulouse.

Milgrom, P. & J. Roberts (1991). "Adaptative and sophisticated learning in repeated normal-form games", *Games and Economic Behavior* **3**, 82-100.

Nachbar, J. (1990): "Evolutionary selection in dynamic games", *International Journal of Game Theory* **19**, 59-90.

Pearce, D. (1984): "Rationalizable strategic behavior and the problem of perfection", *Econometrica* **52**, 1029-1050.

Piccione, M. and A. Rubinstein (1993): "Finite automata play a repeated extensive game", *Journal of Economic Theory* **61**, 160-68.

Sabourian, H. (1991), "The folk-theorem of repeated games with bounded (one-period) memory", mimeo.

Sabourian, H., (1997), “Repeated games with M-period bounded memory (pure strategies)”, *Journal of Mathematical Economics*, forthcoming.

Wallner, K., (1999), Sequential moves and tacit collusion: Reaction function cycles in finite pricing duopoly, *Journal of Economic Theory* **84**, 251-267.