

Folk Theorems with Bounded Recall under (Almost) Perfect Monitoring—Addendum*

George J. Mailath[†] Wojciech Olszewski[‡]

Theorem 3 (Barlo, Carmona, and Sabourian (2008)) *Suppose $n = 2$. For all $v \in \mathcal{F}^{\dagger p}$, for all $\varepsilon > 0$, there exists $\bar{\delta} \in (0, 1)$ and a stationary bounded recall strategy profile that is, for all $\delta \in (\bar{\delta}, 1)$, a patiently strict subgame perfect equilibrium with discounted average payoff within ε of v .*

Proof. Recall that \hat{a}_{-i}^i denotes the action of player $-i$ that minmaxes player $i = 1, 2$; to simplify notation, we write \hat{a} for mutual minmax $(\hat{a}_1^2, \hat{a}_2^1)$. Let \hat{b}_{-i} denote an action of player $-i$ distinct from \hat{a}_{-i}^i , and set $\hat{b} \equiv (\hat{b}_1, \hat{b}_2)$. There is a cycle of action profiles $h^n \equiv (a^1, \dots, a^n)$ whose average payoff is within $\varepsilon/2$ of v . Without loss of generality assume that:

1. the first $k \in \{2, \dots, n-1\}$ action profiles of the cycle are \hat{b} ,
2. none of the remaining $n-k$ action profiles is \hat{b} , and
3. the $(k+1)$ -st action profile of the cycle is \hat{a} .

Suppose that $T-n$ periods of playing the mutual minmax suffice to make any sequence of n unilateral deviations unprofitable; more precisely, suppose that

$$n \cdot \left[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a) \right] < (T-n) \cdot [v_i - u_i(\hat{a})], \quad \forall i = 1, 2. \quad (1)$$

According to the strategy profile to be defined, players are prescribed to play the next action profile in the cycle h^n if the last ℓ periods of the history is consistent with the first ℓ periods of the cycle, for $\ell = k, k+1, \dots, n-1$.

*January 21, 2011. See Mailath and Olszewski (2011) for notation.

[†]University of Pennsylvania

[‡]Northwestern University

Similarly, they are prescribed to play \widehat{b} if \widehat{b} had been played in the $\ell = 0, \dots, k-1$ most recent periods (i.e., the first ℓ periods of the cycle has been played), and in the immediately preceding periods, either a full cycle h^n or T consecutive periods of \widehat{a} had been played. In all other cases, players are prescribed to play \widehat{a} . By construction, the prescribed repeated-game strategies have bounded recall of length $T + k - 1$.

The profile is described by an automaton with states $\{w(0, \ell) : \ell = 1, \dots, n\} \cup \{w^*(0, 1)\} \cup \{w(1, \ell) : \ell = 1, \dots, T\} \cup \{w(2, \ell) : \ell = 2, \dots, k\}$, initial period $w^0 = w(0, 1)$, and output function $f(w^*(0, 1)) = \widehat{b}$, $f(w(0, \ell)) = a^\ell$, and $f(w(1, \ell)) = f(w(2, \ell)) = \widehat{a}$ for all ℓ . For states $w(0, \ell)$, transitions are given by

$$\tau(w(0, \ell), a) = \begin{cases} w(0, \ell + 1), & \text{if } \ell \leq n - 1 \text{ and } a = a^\ell, \\ w(0, 1), & \text{if } \ell = n \text{ and } a = a^\ell, \\ w(1, m + 1), & \text{if } a = \widehat{a} \neq a^\ell, \\ w(1, 1), & \text{if } \ell \leq n \text{ and } a \neq a^\ell, \widehat{a}, \widehat{b}; \end{cases}$$

where if $\ell \geq 2$, $m - 1$ is the number of consecutive action profiles \widehat{a} that immediately precede a^ℓ in h^n (i.e., $a^{\ell-(m-1)} = \dots = a^{\ell-1} = \widehat{a}$), and if $\ell = 1$, $m - 1$ is the number of consecutive action profiles \widehat{a} that end h^n (i.e., $a^{n-(m-2)} = \dots = a^n = \widehat{a}$, and $m = 1$ if $a^n \neq \widehat{a}$);

$$\tau(w^*(0, 1), a) = \begin{cases} w(0, 2), & \text{if } a = \widehat{b}, \\ w^*(0, 1), & \text{if } a = \widehat{a} \\ w(1, 1), & \text{if } a \neq \widehat{a}, \widehat{b}, \end{cases}$$

$$\text{and } \tau(w(0, \ell), \widehat{b}) = \begin{cases} w(0, k + 1), & \text{if } \ell = k + 1, \\ w(2, 2), & \text{if } \ell > k + 1. \end{cases}$$

Recall that $a^\ell \neq \widehat{b}$ for $\ell \geq k + 1$, so the first two and the last two cases in the definition of $\tau(w(0, \ell), a)$ are mutually exclusive. Moreover, for $\ell \leq k$, $a^\ell = \widehat{b}$, and so all transitions from $w(0, \ell)$ are described. For states $w(1, \ell)$ and $w(2, \ell)$, the transitions are given by

$$\tau(w(1, \ell), a) = \begin{cases} w(1, \ell + 1), & \text{if } \ell \leq T - 1 \text{ and } a = \widehat{a}, \\ w^*(0, 1), & \text{if } \ell = T \text{ and } a = \widehat{a}, \\ w(2, 2), & \text{if } \ell \leq T \text{ and } a = \widehat{b}, \text{ and} \\ w(1, 1), & \text{if } \ell \leq T \text{ and } a \neq \widehat{a}, \widehat{b}, \end{cases}$$

and

$$\tau(w(2, \ell), a) = \begin{cases} w(2, \ell + 1), & \text{if } \ell \leq k - 1 \text{ and } a = \widehat{b}, \\ w(0, k + 1), & \text{if } \ell = k \text{ and } a = \widehat{b}, \\ w(1, 2), & \text{if } \ell \leq k \text{ and } a = \widehat{a}, \text{ and} \\ w(1, 1), & \text{if } \ell \leq k \text{ and } a \neq \widehat{a}, \widehat{b}. \end{cases}$$

The “bad” states $w(1, \ell)$ and $w(2, \ell)$ encode the number of consecutive action profiles \widehat{a} and \widehat{b} , respectively, that precede the current period. That is, $w(1, \ell)$ means that \widehat{a} has been played in the $\ell - 1$ most recent periods, and $w(2, \ell)$ means that \widehat{b} has been played in the $\ell - 1$ most recent periods. If action profile \widehat{b} happens to be played in k consecutive periods, then the profile *reinitializes*: players are prescribed to continue playing cycle h^n , independent of the history more than k periods ago. It is important that $k \geq 2$: If k were equal to 1, a unilateral deviation in states $w(0, \ell)$, $\ell > k$, could cause a (possibly profitable) transition to state $w(0, k + 1)$.

We now verify that the profile is a strict subgame perfect equilibrium. By a unilateral deviation in state $w(0, \ell)$, a player gains at most $n \cdot [\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a)]$. Indeed, the player may gain not only in terms of the current flow payoff, but also in terms of the flow payoffs in other periods remaining to the end of the cycle, because for some periods of the cycle, the payoff to playing \widehat{a} may exceed the payoff to playing the prescribed strategy. However, the gain can be estimated by $[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a)]$ times the number of remaining periods.

The player loses (approximately) $v_i - u_i(\widehat{a})$ in each of the next periods of mutual minmax, after the cycle ends. And there are at least $T - n$ such periods. Indeed, any sequence of consecutive action profiles \widehat{a} at the end of the cycle counts to the T periods of mutual minmax following the deviation, but such a sequence cannot be longer than n (or more precisely, $n - 1$), because a unilateral deviation in the first k periods of the cycle cannot yield action profile \widehat{a} . Thus, any unilateral deviation is unprofitable by virtue of (1).

Consider now any state $w(1, \ell)$. Players are prescribed to play \widehat{a} . Note that no unilateral deviation can result in $a = \widehat{b}$. If players play the prescribed action profile \widehat{a} , the next state will be $w(1, \ell + 1)$ (in the case when $\ell \leq T - 1$) or $w(0, 1)$ (when $\ell = T$). After a unilateral deviation, the next state will be $w(1, 1)$. Players prefer to play the prescribed strategies, since this results in $T - \ell + 1$ periods of mutual minmax, followed by playing the cycle. Any unilateral deviation results instead in (at best) the minmax payoff in the current period, followed by T periods of mutual minmax and then playing the cycle. The former option dominates the latter for sufficiently large δ .

Similarly in every state $w(2, \ell)$, players are prescribed to play \hat{a} , and the next state will be $w(1, 2)$ or $w(1, 1)$, depending on whether the prescribed action profile $a = \hat{a}$ or a unilateral deviation occurs. Players prefer to play the prescribed strategies, since this results in T periods of mutual minmax, followed by playing the cycle. Any unilateral deviation results instead in (at best) the minmax payoff in the current period, followed by T periods of mutual minmax and then playing the cycle.

Proving patient strictness is similar to the arguments in the proof of Lemma 2. ■

References

- BARLO, M., G. CARMONA, AND H. SABOURIAN (2008): “Bounded Memory with Finite Action Spaces,” Sabanci University, Universidade Nova de Lisboa and University of Cambridge.
- MAILATH, G. J., AND W. OLSZEWSKI (2011): “Folk Theorems with Bounded Recall under (Almost) Perfect Monitoring,” *Games and Economic Behavior*, 71(1), 174–192.