The Value of Recall

Ron Peretz*

March 25, 2007

1 Extended Abstract

This work studies two-person zero-sum repeated games in which at least one of the players is restricted to (mixtures of) bounded recall strategies. A (pure) k recall strategy is a strategy that relies only on the last k periods of history. This work improves previous results [2, 4] on repeated games with bounded recall by extending the range of settings for which we can approximate the value.

Bounded recall is one of the alternatives proposed by Aumann [1] to model limited rationality in repeated games. Lehrer [2] studied infinitely repeated two-player zero-sum games where both players have bounded recall. Neyman and Okada [3, 4] study a setting in which one player is bounded while the other is fully rational. In [4] they examine specifically the case of bounded recall. The current work extends results of both [2] and [4].

Apparently, in all previous works, the limited player secures the minimax payoff by playing an oblivious[1] strategy. Our main result follows this line:[2]

Theorem 1.1. *Let $G =< I, J, g >$ be a two-player zero-sum game in strategic normal form. For every sequence of positive integers $\{T_k\}_{k=1}^{\infty}$ and every $h \neq \log_2 I$, if $\lim_{k\to\infty} \frac{\log_2 T_k}{k} = h$, then*[3]

$$\begin{aligned} \lim_{k\to\infty} \mathrm{val}G^{T_k}(k_{obl}, \infty) &= \\ \lim_{k\to\infty} \mathrm{val}G^{\infty}(k_{obl}, T_k) &= \max_{\substack{\sigma\in\Delta(I):\\ H(\sigma)\geq h, \text{or}\\ H(\sigma)=0}} \min_{\tau\in J} G(\sigma,\tau). \end{aligned}$$

*Institute of Mathematics, The Hebrew University of Jerusalem. Research supported in part by Israel Science Foundation grant 263/03. This work is the author's Master's thesis (2005). The author would like to thank his advisor Prof. Abraham Neyman for his close guidance and endless patience.
Email: ronprtz@math.huji.ac.il.

[1] "Oblivious" means that the actions taken by the player do not depend on the history of actions of the other player, but only on his own.

[2] See theorem 1.1: the former result (regarding G^{T_k}) extends [4]. The later extends [2].

[3] $G^{\alpha}(k_{obl}, m)$ is the α fold repeated version of G where player one is restricted to oblivious k-recall strategies and player two is restricted to (non-oblivious) m-recall strategies.

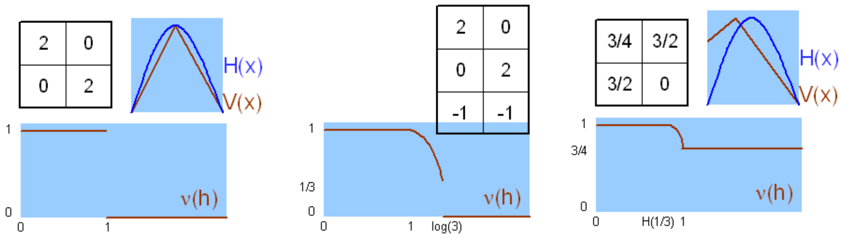

Figure 1: examples (left to right) – “matching pennies”, “matching pennies+”, and a game with a continuous ν.

1.1 Examples

Let us denote

$$\begin{aligned}
\nu_* &:= \max_{i \in I} \min_{j \in J} G(i,j) \\
\tilde{\nu}(h) &:= \max_{\substack{\sigma \in \Delta(I): \\ H(\sigma) \geq h}} \min_{\tau \in J} G(\sigma,\tau) \quad = \min_{\tau \in \Delta(J)} \max_{\substack{\sigma \in \Delta(I): \\ H(\sigma) \geq h}} G(\sigma,\tau) \\
\nu(h) &:= \max_{\substack{\sigma \in \Delta(I): \\ H(\sigma) \geq h, \text{or} \\ H(\sigma)=0}} \min_{\tau \in J} G(\sigma,\tau) = \tilde{\nu}(h) \vee \nu_*.
\end{aligned}$$

Consider the game of “matching pennies” described in figure 1. Since the optimal strategy in this game is $(\frac{1}{2}, \frac{1}{2})$ which is also the one with maximal entropy, the theorem, roughly, says that if $\frac{\log_2 T_k}{k} < 1$ then the value of the repeated game (in either one of the settings), ν, is “equal” to the value of the one stage game. Since the inferior player cannot expect anything greater than the value of the stage game, the value that can be obtained by an oblivious agent matches the value that can be obtained by a non-oblivious agent. Whether this is the case in general seems unlikely, yet it is unknown[4].

The function $\nu(h)$ is continuous at all but maybe one point $h = \log_2(I)$. In the above example, ν is *not* continuous at that suspicious point, $h = 1$. It can be shown that $\lim_{k\to\infty} \frac{\log_2 T_k}{k} = h$ implies the convergence of the value of the repeated games if and only if ν is continuous at h. The third example in figure 1 is a game for which ν *is* continuous at the suspicious point (and any other point).

Finally, let us look at the game – “matching pennies+”. The third alternative of player one is strongly dominated in the one step game. Nevertheless, in the repeated game, player one can gain from playing the third alternative occasionally. An intuitive explanation is that the “memory” of player one is in his actions, and more memory means longer endurance along the repeated game. By playing the third alternative here-and-there he might lose a few rounds, but gain more memory, endure longer, and thus improve the overall payoff.

[4] It is known to be true for any game in which player one – the inferior player – has only two (pure) alternatives. Namely, $|I| = 2$. More generally, the value of the repeated game in the non-oblivious setting is asymptotically $\leq \min_{\tau \in \Delta(J)} \max_{\substack{\sigma \in \Delta(I): \\ H(\sigma) \geq h - H(\tau)}} G(\sigma,\tau) \vee \nu_*$.

1.2 Outlines of the Proof

The last step of the proof and the first step towards is the following lemma that suggests how to design efficient oblivious strategies:[5]

Lemma (Neyman-Okada's Criterion)**.** *Let $p \in \Delta(I)$ be a mixed strategy in the one stage game that secures a payoff v. In order for a player to secure a payoff of $v - \epsilon$ in the T-stage repeated game, it is sufficient for him to be able to implement an I^T-valued random variable x that satisfies: (i) The average entropy per stage is close to the entropy of p, namely $\frac{1}{T}H(x) \geq H(p) - \delta$; and (ii) the mean empirical distribution is close to p, namely $\|\mathbf{E}[\mathrm{emp}(x)] - p\| < \delta$.*

The objective, therefore, is to find conditions on T, k and p that will ensure the existence of a *large* number (satisfying property (i) of the lemma) of sequences of length T, with empirical distribution close to p (satisfying (ii)) which are implementable by (oblivious) k-recall strategies. To start with, the reader may consider the case of p being the uniform distribution, in which case the question is "how many sequences of length T can be implemented by k-recall strategies?". Since an asymptotic answer is sufficient, it turns out that the problem reduces to approximating the number of sequences that do not contain any combination of k consecutive elements more than once. A partial answer, yet sufficient for deriving theorem 1.1, is the following theorem:

Theorem 1.2. *Let $\{T_k\}_{k=1}^{\infty}$ be a sequence of positive integers, $T_k \to \infty$. Let $x_1, x_2, \ldots$ be a sequence of i.i.d. random variables with common distribution p (over a finite set of values). If $\limsup \frac{\log_2 T_k}{k} < H(p)$, then*

$$\mathbf{P}\big(\forall 0 \leq t \neq s \leq T_k \quad (x_{t+1}, \ldots, x_{t+k}) \neq (x_{s+1}, \ldots, x_{s+k})\big) = \exp(-\mathrm{o}(T_k)).$$

The assumption that $\limsup \frac{\log_2 T_k}{k} < H(p)$ is necessary[6]. On the other hand, the theorem does not tell us how small $\mathrm{o}(T_k)$ is. It is of interest – finding an explicit expression for (the asymptotic of) the probability above.

References

[1] Aumann R.J. (1981). "Survey of Repeated Games," *Essays in Game Theory and Mathematical Economics in Honor of Oscar Morgenstern*, 11-42. Mannheim: Bibliographisches Institut.

[2] Lehrer, E. (1994). "Finiely Many Players with Bounded Recall in Infinitely Repeated Games," *Games Econom. Behav.* **7**, 390-405.

[3] Neyman, A., and Okada, D. (1997). "Repeated Games with Bounded Entropy," *Games Econom. Behav.* **30**, 228-247.

[4] Neyman, A., and Okada, D. (2005). "Growth of Strategy Sets, Entropy, and Nonstationary Bounded Recall," DP 411, Center for Rationality, Hebrew University.

[5] The original statement is more general. Here it is applied to the setting of $G^T(k_{obl}, \infty)$, and the quantifiers around ϵ and δ were omitted for readability.

[6] Considering only s and t congruent to 1 modulo k one obtains an instance of the well studied "birthday" problem.