The Optimality of Regret Matching

Sergiu Hart

July 2008

THE OPTIMALITY OF REGRET MATCHING

Sergiu Hart

Center for the Study of Rationality
Dept of Economics   Dept of Mathematics
The Hebrew University of Jerusalem

hart@huji.ac.il
http://www.ma.huji.ac.il/hart

Joint work with

Elchanan Ben-Porath

Repeated Games

In a repeated game:

Repeated Games

In a repeated game:

- If the opponent plays ***i.i.d.*** (independently identically distributed)

Repeated Games

In a repeated game:

- If the opponent plays ***i.i.d.*** (independently identically distributed)

  Then ***fictitious play*** is optimal

Repeated Games

In a repeated game:

- If the opponent plays ***i.i.d.*** (independently identically distributed)

  Then ***fictitious play*** is optimal

- If the opponent is **OBLIVIOUS**

Repeated Games

In a repeated game:

- If the opponent plays ***i.i.d.*** (independently identically distributed)

  Then ***fictitious play*** is optimal

- If the opponent is **OBLIVIOUS**

OBLIVIOUS = does not see / know / use my actions

Repeated Games

In a repeated game:

- If the opponent plays ***i.i.d.*** (independently identically distributed)

  Then ***fictitious play*** is optimal

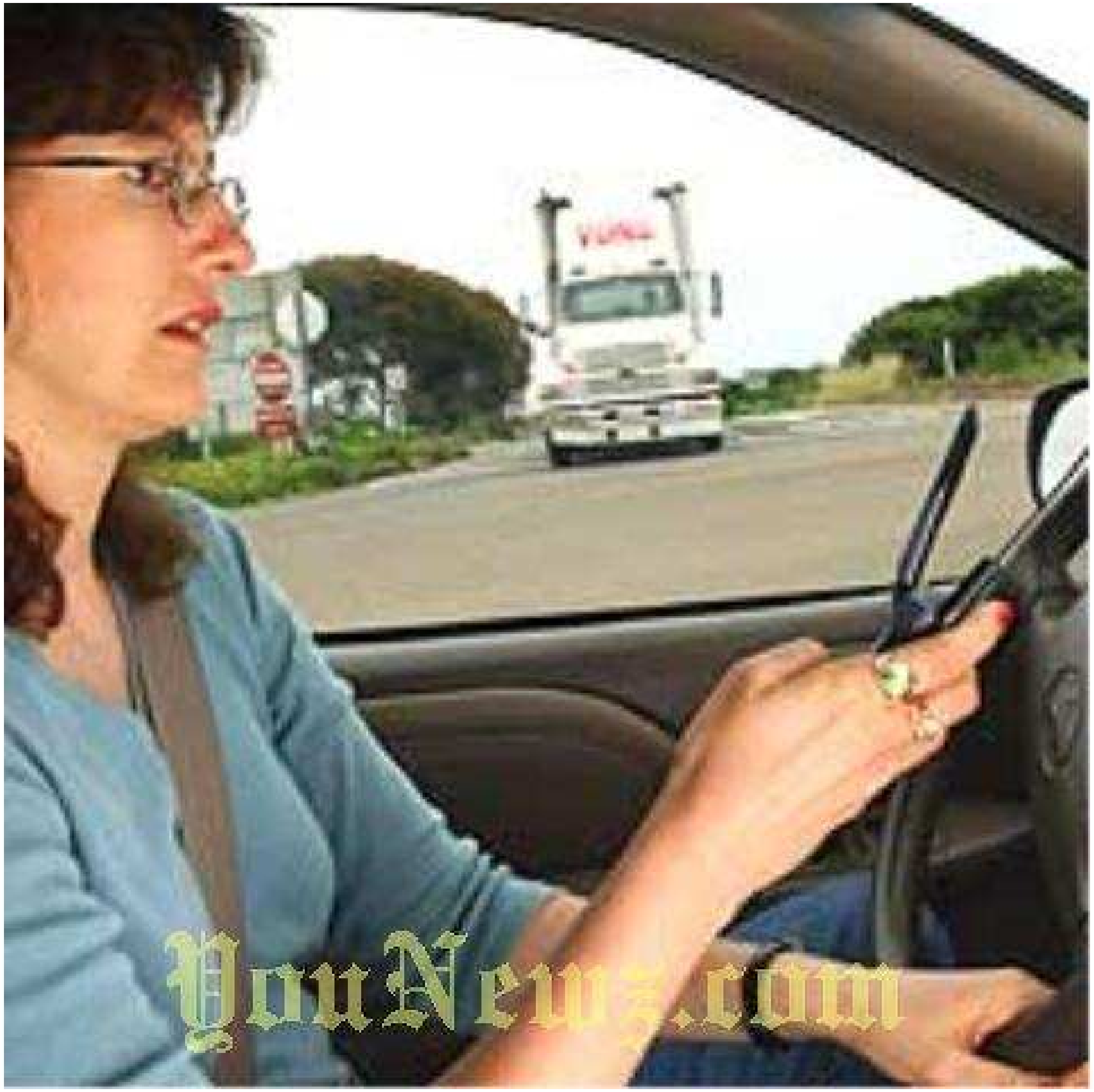
- If the opponent is **OBLIVIOUS**

  Then ?

OBLIVIOUS = does not see / know / use my actions

Oblivious

Oblivious



Model: One-Shot Game

Model: One-Shot Game

The ***one-shot game*** $\Gamma = < I, J, u >$

Model: One-Shot Game

The ***one-shot game*** $\Gamma = < I, J, u >$

- I = (finite) set of *actions* of Player 1
- J = (finite) set of *actions* of Player 2 (the "opponent")

Model: One-Shot Game

The ***one-shot game*** $\Gamma = < \boldsymbol{I}, \boldsymbol{J}, \boldsymbol{u} >$

- $\boldsymbol{I}$ = (finite) set of *actions* of Player 1
- $\boldsymbol{J}$ = (finite) set of *actions* of Player 2 (the "opponent")
- $\boldsymbol{u} : \boldsymbol{I} \times \boldsymbol{J} \longrightarrow \mathbb{R}$
  = the *payoff function* of Player 1

Model: One-Shot Game

The ***one-shot game*** $\Gamma = < \boldsymbol{I}, \boldsymbol{J}, \boldsymbol{u} >$

- $\boldsymbol{I}$ = (finite) set of *actions* of Player 1
- $\boldsymbol{J}$ = (finite) set of *actions* of Player 2 (the "opponent")
- $\boldsymbol{u} : \boldsymbol{I} \times \boldsymbol{J} \longrightarrow \mathbb{R}$
  = the *payoff function* of Player 1
  - without loss of generality assume:
    $\mathbf{0} \leq \boldsymbol{u}(\boldsymbol{i}, \boldsymbol{j}) \leq \mathbf{1}$ for all $\boldsymbol{i} \in \boldsymbol{I}, \boldsymbol{j} \in \boldsymbol{J}$

Model: Repeated Game

Model: Repeated Game

The ***repeated game*** Γ^{∞}

Model: Repeated Game

The *repeated game* $\Gamma_{\text{O}}^{\infty}$

- Player 2 is restricted to OBLIVIOUS strategies $\eta : \cup_{t=0}^{\infty} J^t \longrightarrow \Delta(J)$

Model: Repeated Game

The *repeated game* $\Gamma_{\mathbf{o}}^{\infty}$

- Player 2 is restricted to **OBLIVIOUS** strategies
  $\eta : \cup_{t=0}^{\infty} J^t \longrightarrow \Delta(J)$
- Player 1 is unrestricted
  $\sigma : \cup_{t=0}^{\infty} (I \times J)^t \longrightarrow \Delta(I)$

Model: Repeated Game

The *repeated game* $\Gamma_{\text{o}}^{\infty}$

- Player 2 is restricted to OBLIVIOUS strategies
  $\eta : \cup_{t=0}^{\infty} J^t \to \Delta(J)$
- Player 1 is unrestricted
  $\sigma : \cup_{t=0}^{\infty} (I \times J)^t \to \Delta(I)$
- *Payoffs*
  - $v_t := u(i_t, j_t)$
    payoff at time t
  - $\bar{v}_T := (1/T) \sum_{t=1}^{T} v_t$
    average payoff up to time T

Model: Repeated Game

The ***repeated game*** $\Gamma_{\mathbf{o}}^{\infty}$

- Player 2 is restricted to OBLIVIOUS strategies
  $\eta : \cup_{t=0}^{\infty} J^t \longrightarrow \Delta(J)$
- Player 1 is unrestricted
  $\sigma : \cup_{t=0}^{\infty} (I \times J)^t \longrightarrow \Delta(I)$
- Without loss of generality:
  Player 1 uses "self-oblivious" strategies
  $\sigma : \cup_{t=0}^{\infty} J^t \longrightarrow \Delta(I)$

Dominance

Dominance

- If Player 1 has a *dominant* action $i \in I$ in Γ then playing i forever is a dominant strategy in the repeated game

Dominance

- If Player 1 has a *dominant* action $i \in I$ in Γ then playing i forever is a dominant strategy in the repeated game

- The one-shot game Γ is called ***essential***

Dominance

- If Player 1 has a *dominant* action $i \in I$ in Γ then playing i forever is a dominant strategy in the repeated game

- The one-shot game Γ is called ***essential*** if Player 1 does **NOT** have a dominant action:

Dominance

- If Player 1 has a *dominant* action $i \in I$ in Γ then playing i forever is a dominant strategy in the repeated game

- The one-shot game Γ is called ***essential*** if Player 1 does **NOT** have a dominant action:

  for every $i \in I$ there exists $i' \in I$ with $u(i, j) < u(i', j)$ for some $j \in J$

Dominance

Proposition. *Let* Γ *be an essential game.*

Dominance in Γ_o^∞

Proposition. *Let* Γ *be an essential game. Then every strategy of Player 1 in* Γ_o^∞ *is* **weakly dominated**.

Dominance in Γ_o^∞

Proposition. *Let Γ be an essential game. Then every strategy of Player 1 in Γ_o^∞ is* **weakly dominated***.*

That is, for every σ there exists $\hat{\sigma}$ such that:

- For every oblivious strategy η of Player 2

$$\mathbf{E}_{\hat{\sigma},\eta}\left[\bar{v}_T\right] \geq \mathbf{E}_{\sigma,\eta}\left[\bar{v}_T\right] - 1/T \quad \text{for all } T \geq 1$$

Dominance in $\Gamma_{\text{o}}^{\infty}$

Proposition. *Let* Γ *be an essential game. Then every strategy of Player 1 in* $\Gamma_{\text{o}}^{\infty}$ *is* **weakly dominated**.

That is, for every σ there exists $\hat{\sigma}$ such that:

- For every oblivious strategy η of Player 2

$$\mathbf{E}_{\hat{\sigma},\eta}\left[\bar{v}_T\right] \geq \mathbf{E}_{\sigma,\eta}\left[\bar{v}_T\right] - 1/T \quad \text{for all } T \geq 1$$

- There exists an oblivious strategy η_0 of Player 2 and a constant $\gamma > 0$ such that

$$\mathbf{E}_{\hat{\sigma},\eta_0}\left[\bar{v}_T\right] > \mathbf{E}_{\sigma,\eta_0}\left[\bar{v}_T\right] + \gamma \qquad \text{for all } T \geq 1$$

Regret Matching

Regret Matching

For every $\boldsymbol{k} \in \boldsymbol{K}$, let $\boldsymbol{\sigma_k} : \cup_{\boldsymbol{t=0}}^{\infty} \boldsymbol{J^t} \rightarrow \boldsymbol{\Delta(I)}$ be a (self-oblivious, behavior) strategy of Player 1.

Regret Matching

For every $k \in K$, let $\sigma_k : \cup_{t=0}^{\infty} J^t \to \Delta(I)$ be a (self-oblivious, behavior) strategy of Player 1.

Theorem. *For every finite set K there exists a K-REGRET-MATCHING strategy $\sigma^* \equiv \sigma^*_K$ such that*

$$\mathbf{E}_{\sigma^*,\eta}\left[\bar{v}_T\right] \geq \mathbf{E}_{\sigma_k,\eta}\left[\bar{v}_T\right] - \frac{3\sqrt{|K|}}{\sqrt{T}}$$

for every $k \in K$, every time $T \geq 1$, and every oblivious strategy η of Player 2.

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows.

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$
    (the total payoff of σ_k against the realized sequence of actions of Player 2)

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$ (the total payoff of σ_k against the realized sequence of actions of Player 2)
  - $R(k) := [V(k) - U]_+$ (the **REGRET** of k)

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$ (the total payoff of σ_k against the realized sequence of actions of Player 2)
  - $R(k) := [V(k) - U]_+$ (the **REGRET** of k)
  - $\rho_k := R(k) / \sum_{\ell \in K} R(\ell)$ (the normalized regret of k)

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$
  - $R(k) := [V(k) - U]_+$ (the **REGRET** of k)
  - $\rho_k := R(k) / \sum_{\ell \in K} R(\ell)$

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$
  - $R(k) := [V(k) - U]_+$ (the **REGRET** of k)
  - $\rho_k := R(k) / \sum_{\ell \in K} R(\ell)$
- σ^* plays like σ_k with probability ρ_k

Regret Matching

The K-**REGRET-MATCHING** strategy $\sigma^* \equiv \sigma^*_K$ is defined as follows. At time $T+1$:

- $U := \sum_{t=1}^{T} v_t$ (realized total payoff)
- For each $k \in K$:
  - $V(k) := \sum_{t=1}^{T} u(\sigma_k(h^2_{t-1}), j_t)$
  - $R(k) := [V(k) - U]_+$ (the **REGRET** of k)
  - $\rho_k := R(k) / \sum_{\ell \in K} R(\ell)$
- σ^* plays like σ_k with probability ρ_k :

$$\sigma^*(h^2_T)(i) = \sum_{k \in K} \rho_k \cdot \sigma_k(h^2_T)(i)$$

Regret Matching

Regret Matching

- Generalizes Hannan's result:

Regret Matching

- Generalizes Hannan's result:
  - $\sigma_i =$ "play i forever"
    for each $i \in I$

Regret Matching

- Generalizes Hannan's result:
  - $\boldsymbol{\sigma}_{\boldsymbol{i}} = $ "play $\boldsymbol{i}$ forever" for each $\boldsymbol{i} \in \boldsymbol{I}$
- Blackwell approachability (with changing payoffs)

Regret Matching

Regret Matching - Countable

Theorem. *For every* COUNTABLE SET K *there exists a* K-REGRET-MATCHING *strategy* $\sigma^* \equiv \sigma^*_K$ *such that*

Regret Matching - Countable

Theorem. *For every* COUNTABLE SET K *there exists a* K-REGRET-MATCHING *strategy* $\sigma^* \equiv \sigma^*_K$ *such that*

$$\mathbf{E}_{\sigma^*,\eta}\left[\bar{v}_T\right] \geq \mathbf{E}_{\sigma_k,\eta}\left[\bar{v}_T\right] - O\left(T^{-1/4}\right)$$

for every $k \in K$*, every time* $T \geq 1$*, and every oblivious strategy* η *of Player 2.*

Regret Matching - Countable

Regret Matching - Countable

The K-REGRET-MATCHING strategy $\sigma^* \equiv \sigma^*_K$ for COUNTABLE $K = 1, 2, \ldots$:

Regret Matching - Countable

The K-REGRET-MATCHING strategy $\sigma^* \equiv \sigma^*_K$ for COUNTABLE $K = 1, 2, \ldots$:

- For each $k = 1, 2, \ldots$:
  - BLOCK k has $n_k = k^3$ periods

Regret Matching - Countable

The K-REGRET-MATCHING strategy $\sigma^* \equiv \sigma^*_K$ for COUNTABLE $K = 1, 2, ...$:

- For each $k = 1, 2, ...$:
  - BLOCK k has $n_k = k^3$ periods
  - During BLOCK k play the regret-matching strategy for the set $\{\sigma_1, ..., \sigma_k\}$

Regret Matching - Countable

The $\boldsymbol{K}$-REGRET-MATCHING strategy $\boldsymbol{\sigma^* \equiv \sigma^*_K}$ for COUNTABLE $\boldsymbol{K = 1, 2, \ldots}$:

- For each $\boldsymbol{k = 1, 2, \ldots}$:
  - BLOCK $\boldsymbol{k}$ has $\boldsymbol{n_k = k^3}$ periods
  - During BLOCK $\boldsymbol{k}$ play the regret-matching strategy for the set $\boldsymbol{\{\sigma_1, \ldots, \sigma_k\}}$
    - Regrets are computed on the basis of the *current block* only

Regret Matching

Regret Matching

Countable sets of strategies:

Regret Matching

Countable sets of strategies:

- Automata

Regret Matching

Countable sets of strategies:

- Automata
- Turing Machines

Regret Matching

Countable sets of strategies:

- Automata
- Turing Machines
- A **COMPUTABLE SET OF STRATEGIES**

Regret Matching

Countable sets of strategies:

- Automata
- Turing Machines
- A **COMPUTABLE SET OF STRATEGIES**
  - **INPUT**:
    - k (index of a strategy)
    - h_t (a history)

Regret Matching

Countable sets of strategies:

- Automata
- Turing Machines
- A **COMPUTABLE SET OF STRATEGIES**
  - **INPUT**:
    - k (index of a strategy)
    - h_t (a history)
  - **OUTPUT**:
    - $\sigma_k(h_t) \in \Delta(I)$

Regret

Regret: The End