Adaptive Heuristics

Game Theory Presentation

Théo Delemazure

Ecole normale supérieure, PSL University

March 23, 2020

Paper presentation

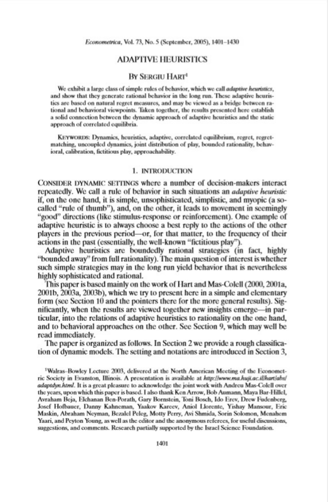
Title: 8L9q§_®0 [0«}_ §_H

Author: ~0}X_« [9}§ (Former president of the W9h0 ¦ \0n}´ ~nH_0§´)

Year: ©··U

Journal: NHnjnh0§}_H9

Citations: 279 to this day



Econometrica, Vol. 73, No. 5 (September, 2005), 1401–1430

ADAPTIVE HEURISTICS

BY SERGIU HART[1]

We exhibit a large class of simple rules of behavior, which we call *adaptive heuristics*, and show that they generate rational behavior in the long run. These adaptive heuristics are based on natural regret measures, and may be viewed as a bridge between rational and behavioral viewpoints. Taken together, the results presented here establish a solid connection between the dynamic approach of adaptive heuristics and the static approach of correlated equilibria.

KEYWORDS: Dynamics, heuristics, adaptive, correlated equilibrium, regret, regret-matching, uncoupled dynamics, joint distribution of play, bounded rationality, behavioral, calibration, fictitious play, approachability.

1. INTRODUCTION

CONSIDER DYNAMIC SETTINGS where a number of decision-makers interact repeatedly. We call a rule of behavior in such situations an *adaptive heuristic* if, on the one hand, it is simple, unsophisticated, simplistic, and myopic (a so-called "rule of thumb"), and, on the other, it leads to movement in seemingly "good" directions (like stimulus-response or reinforcement). One example of adaptive heuristic is to always choose a best reply to the actions of the other players in the previous period—or, for that matter, to the frequency of their actions in the past (essentially, the well-known "fictitious play").

Adaptive heuristics are boundedly rational strategies (in fact, highly "bounded away" from full rationality). The main question of interest is whether such simple strategies may in the long run yield behavior that is nevertheless highly sophisticated and rational.

This paper is based mainly on the work of Hart and Mas-Colell (2000, 2001a, 2001b, 2003a, 2003b), which we try to present here in a simple and elementary form (see Section 10 and the pointers there for the more general results). Significantly, when the results are viewed together new insights emerge—in particular, into the relations of adaptive heuristics to rationality on the one hand, and to behavioral approaches on the other. See Section 9, which may well be read immediately.

The paper is organized as follows. In Section 2 we provide a rough classification of dynamic models. The setting and notations are introduced in Section 3,

[1]Walras–Bowley Lecture 2003, delivered at the North American Meeting of the Econometric Society in Evanston, Illinois. A presentation is available at *http://www.ma.huji.ac.il/hart/abs/adaptdyn.html*. It is a great pleasure to acknowledge the joint work with Andreu Mas-Colell over the years, upon which this paper is based. I also thank Ken Arrow, Bob Aumann, Maya Bar-Hillel, Avraham Beja, Elchanan Ben-Porath, Gary Bornstein, Toni Bosch, Ido Erev, Drew Fudenberg, Josef Hofbauer, Danny Kahneman, Yaakov Kareev, Aniol Llorente, Yishay Mansour, Eric Maskin, Abraham Neyman, Bezalel Peleg, Motty Perry, Avi Shmida, Sorin Solomon, Menahem Yaari, and Peyton Young, as well as the editor and the anonymous referees, for useful discussions, suggestions, and comments. Research partially supported by the Israel Science Foundation.

Summary

Summary

Degrees of rationality

We can define the degree of Rationality of a strategy as the complexity of the reasoning and computation for the player using it.

Low rationality ! High rationality

N®ne«§_nj9}´ K´j9h_H	d09}j_jX K´j9h_H
Player , Population of individuals who play an action dictated by their genotype. Two main forces : Selection and Mutation.	Bayesian learning : each player play according to his belief on the world, which is updated every period.

Degrees of rationality

We can define the degree of Rationality of a strategy as the complexity of the reasoning and computation for the player using it.

Low rationality ! High rationality

N®ne«§_nj9}´ K´j9h_H	d09}j_jX K´j9h_H
Player , Population of individuals who play an action dictated by their genotype. Two main forces : Selection and Mutation.	Bayesian learning : each player play according to his belief on the world, which is updated every period.

Degrees of rationality

We can define the degree of Rationality of a strategy as the complexity of the reasoning and computation for the player using it.

Low rationality	!	High rationality
N®ne«§_nj9}´ K´j9h_H		d09}j_jX K´j9h_H
Player , Population of individuals who play an action dictated by their genotype. Two main forces : Selection and Mutation.		Bayesian learning : each player play according to his belief on the world, which is updated every period.

What about medium rationality ?

8L9q§_®0

The player tends to choose better actions every time.

[0«}_ §_H

Define simple, unsophisticated, simplistic rules (Ȳ}«e0 nT §\« h @Ȳ) that the player uses to make his decisions.

What about medium rationality ?

8L9q§_®0

The player tends to choose better actions every time.

[0«}_ §_H

Define simple, unsophisticated, simplistic rules (Ȳ}«e0 nT̄ §\« h@Ȳ) that the player uses to make his decisions.

What about medium rationality ?

8L9q§_®0

The player tends to choose better actions every time.

[0«}_ §_H

Define simple, unsophisticated, simplistic rules (*Ȳ}«e0 nT §\« h @Ȳ*) that the player uses to make his decisions.

Example : Fictitious Play

Ÿ8§ 09H\ q0}_nLJ qℯ9´ §\0 9H§_nj ° _H\ _ §\0 @0 § }0qℯ´ §n 9ℯℯ §\0 q}0®_n«
9H§_nj nT §\0 nqqnj0j§Ÿ

More formally, we have :

$$ {}^{i}(T+1) = argmax_{\ 2A^{i}} \frac{1}{T} \sum_{t=1}^{T} u^{i}(\ ; s_{t}^{\ i}) $$

Notations

§9X0 X9h0	
qℯ9´0}	$1, \dots, N$
i]§«qℯ0 nT 9H§_nj	$S = S_1 \times \dots \times S_N$
q9´n ¸ T«jH§_nj	$u^i : S \to \mathbb{R}$
il̦ n«§	$s^{-i} \in S_1 \times \dots \times S_{i-1} \times S_{i+1} \times \dots \times S_N$
il̦ 9H§_nj	$s^i \in S^i$
g _²0L 9H§_nj	$\sigma^i_t \in \Delta(S^i)$

Perfect monitoring assumption : At the end of each period, all players observe the actions taken by everyone s_t.

Notations

§9X0 X9h0	
qℯ9´0}	$1; \ldots; N$
i]§«qℯ0 nT 9H§_nj	$S = S_1 \ldots S_N$
q9´n¸ T«jH§_nj	$u^i : S \; ! \; \mathbb{R}$
il̜ n«§	$s^{\;i} \; 2 \; S_1 \quad S_{i \; 1} \quad S_{i+1} \quad \ldots \quad S_N$
il̜ 9H§_nj	$s^i \; 2 \; S^i$
g_²0L 9H§_nj	$^i_t \; 2 \quad (S^i)$

Perfect monitoring assumption : At the end of each period, all players observe the actions taken by everyone s_t.

Correlated equilibrium

Before the game starts, each player receives a signal ^{i}. The distribution of signal $= (^1; \ldots ^n)$ is known to all players, and the signal do not affect the payoffs.) It can affect the outcome

(Trivial example : Battle of the Sexes)

Let be a probability distribution on S induced by the probability of signals , then it is a correlated equilibrium iff

$$\forall k \neq j, \sum_{s^{-i} \in S^{-i}} (j; s^{-i}) u^i(j; s^{-i}) \quad \sum_{s^{-i} \in S^{-i}} (j; s^{-i}) u^i(k; s^{-i})$$

Correlated equilibrium

Before the game starts, each player receives a signal ^{i}. The distribution of signal $= (^1, \dots, ^n)$ is known to all players, and the signal do not affect the payoffs. $\Rightarrow$ It can affect the outcome
(Trivial example : Battle of the Sexes)

Let be a probability distribution on S induced by the probability of signals , then it is a correlated equilibrium iff

$$\forall k \neq j, \sum_{s^{-i} \in S^{-i}} (j, s^{-i}) u^i(j, s^{-i}) \quad \sum_{s^{-i} \in S^{-i}} (j, s^{-i}) u^i(k, s^{-i})$$

Correlated equilibrium

Before the game starts, each player receives a signal θ^i. The distribution of signal $\theta = (\theta^1, \dots, \theta^n)$ is known to all players, and the signal do not affect the payoffs. $\Rightarrow$ It can affect the outcome

(Trivial example : Battle of the Sexes)

Let π be a probability distribution on S induced by the probability of signals θ, then it is a correlated equilibrium iff

$$\forall k \neq j, \sum_{s^{-i} \in S^{-i}} \pi(j, s^{-i}) u^i(j, s^{-i}) \geq \sum_{s^{-i} \in S^{-i}} \pi(j, s^{-i}) u^i(k, s^{-i})$$

Correlated equilibrium : The chicken game

	e09®0	§9´
e09®0	5,5	3,6
§9´	6,3	0,0

	e09®0	§9´
e09®0	1/3	1/3
§9´	1/3	0

¦ 9@e0 – The chicken game (left) and its correlated equilibrium (right)

- If a player receive §9´, he knows that the other player received e09®0.
- If a player receives e09®0, then there is a probability of $\frac{1}{2}$ that the other one received §9´ and $\frac{1}{2}$ that he received e09®0. If he follows the recommendation, the expected payoff is 4, and 3 if he choose §9´ instead.

Correlated equilibrium : The chicken game

	e09®0	§9´
e09®0	5,5	3,6
§9´	6,3	0,0

	e09®0	§9´
e09®0	1/3	1/3
§9´	1/3	0

¦9@e0 – The chicken game (left) and its correlated equilibrium (right)

- If a player receive §9´, he knows that the other player received e09®0.
- If a player receives e09®0, then there is a probability of $\frac{1}{2}$ that the other one received §9´ and $\frac{1}{2}$ that he received e09®0. If he follows the recommendation, the expected payoff is 4, and 3 if he choose §9´ instead.

Correlated equilibrium : The chicken game

	e09®0	§9´
e09®0	5,5	3,6
§9´	6,3	0,0

	e09®0	§9´
e09®0	1/3	1/3
§9´	1/3	0

¦ 9@e0 – The chicken game (left) and its correlated equilibrium (right)

- If a player receive §9´, he knows that the other player received e09®0.
- If a player receives e09®0, then there is a probability of $\frac{1}{2}$ that the other one received §9´ and $\frac{1}{2}$ that he received e09®0. If he follows the recommendation, the expected payoff is 4, and 3 if he choose §9´ instead.

Joint distribution of play

Definition

The joint distribution of play is the relative frequency of each N-tuple in the history of play. It is a probability distribution z_T.

$$\forall s \in S; z_T(s) := \frac{1}{T}|\{1 \leq t \leq T : s_t = s\}|$$

It is different from the products of marginals in general.

Summary

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$p_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k \neq j} R_T(j;k) & \text{if } k = j \end{cases}$$

Regret Matching

> Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T} \sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T} \sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$\sigma_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k \neq j} R_T(j;k) & \text{if } k = j \end{cases}$$

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$ {}_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k\neq j} R_T(j;k) & \text{if } k = j \end{cases} $$

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$p^i_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k \neq j} R_T(j;k) & \text{if } k = j \end{cases}$$

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$\sigma_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k\neq j} R_T(j;k) & \text{if } k = j \end{cases}$$

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{\ i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) \quad U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$ {}_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 \quad c\sum_{k\neq j} R_T(j;k) & \text{if } k = j \end{cases} $$

Regret Matching

Switch next period to a different action k with a probability that is proportional to the regret for that action. [5]

- j the action played by i^{th} player at T.
- $U_T := \frac{1}{T}\sum_{t=1}^{T} u^i(s_t)$ the average payoff up to T^{th} period.
- $V_T(j;k) := \frac{1}{T}\sum_{t=1}^{T} v_t^i$ where $v_t^i = u^i(k; s_t^{-i})$ whenever $s_t^i = j$.
- $D_T(j;k) := V_T(j;k) - U_T$ the internal regret associated to action k.
- $R_T(j;k) := [D_T(j;k)]_+$ the non-negative regret.
- The mixed action when j was the last action played is :

$$\sigma_{T+1}(k) = \begin{cases} cR_T(j;k) & \text{if } k \neq j \\ 1 - c\sum_{k \neq j} R_T(j;k) & \text{if } k = j \end{cases}$$

Regret Matching Theorem

Theorem (Hart and Mas-Colell, 2000 [5])

d0§ 09H\ qe9´0} qe9´ }0X}0§ h9§H_jXu ¦ \0j §\0 an_j§ L_ §}_@«§_nj nT qe9´ Hnj®0}X0 §n §\0 0§ nT Hn}}0e9§0L 0x«_e_@}_9 nT §\0 §9X0 X9h0u

Regret Matching Theorem : Proof

Theorem (Hart and Mas-Colell, 2000 [5])

dO§ O9H\ qe9´O} qe9´ }OX}O§ h9§H_jXu ¦ \Oj §\O an_j§ L_ §}_@«§_nj nT qe9´ Hnj®O}XO §n §\O O§ nT Hn}}Oe9§OL Ox«_e_@}_9 nT §\O §9XO X9hOu

) The proof is a variation of *?e9Hc°Oee[9qq}n9H\9@_e_§´ §\On}Oh*.

We denote $_t$ the distance between internal regrets and the negative orthant :

$$_t = [dist(D_t; \mathrm{R}\ )]^2 = \quad D_t \quad D_t \quad ^2 = \quad D_t^+ \quad ^2 = \sum_{k \neq j} D_t^+(k;j)^2$$

We also denote A_t the regret of period t : $D_t = \frac{1}{t} \sum_{i=1}^{T} A_i$.

Regret Matching Theorem : Proof

Step 1/8 : Recursion equation

- Since $D_t \; 2 \; \mathrm{R}$, $_{t+v} \quad D_{t+v} \quad D_t \quad ^2$ and

$$_{t+v} \qquad \frac{1}{t+v}(tD_t + \sum_{w=1} A_{t+w}) \quad D_t \qquad ^2$$

- We have $jA_tj \quad \sum_{k \neq j} jA_t(j;k)j \quad m(m \quad 1)2ju^ij \quad m(m \quad 1)2M$. This gives us a bound on jD_tj.
- With some calculus we obtain :

$$_{t+v} \qquad \frac{t^2}{(t+v)^2} \; _t + \frac{2t}{(t+v)^2} \sum_{w=1} A_{t+w} \quad R_t + \frac{v^2}{(t+v)^2} C$$

with $C = m(m \quad 1)16M^2$.

Regret Matching Theorem : Proof

Step 1/8 : Recursion equation

- This gives us our basic recursion equation :

$$\mathrm{E}[(t+v)^2 \, \rho_{t+v} | h_t] \leq t^2 \rho_t + 2t \sum_{w=1}^{v} R_t \cdot \mathrm{E}[A_{t+w} | h_t] + O(v^2) \qquad (1)$$

- The middle term on right-hand side does not immediately vanish, so we need to estimate this term.

Regret Matching Theorem : Proof

Step 2/8 : Rewriting the middle term

- Using the definition of expectation

$$E[A_{t+w}(j;k)|h_t] = \sum_{s^{\ i}} P((j;s^{\ i})_{t+w}|h_t)[u^i(k;s^{\ i}) \quad u^i(j;s^{\ i})]$$

- Using the definition of the mixed action ${}_t$ and doing some simple calculus, we obtain :

$$R_t \quad E[A_{t+w}|h_t] = \frac{1}{c} \sum_{s^{\ i}} \sum_{j \in S^i} u^i(j;s^{\ i}) \ {}_{t;w}(j;s^{\ i}) \qquad (2)$$

where ${}_{t;w} = \sum_{k \in S^i} {}_t(k \to j) P((k;s^{\ i})_{t+w}|h_t) \quad P((j;s^{\ i})_{t+w}|h_t)$

Regret Matching Theorem : Proof

Step 3/8 : An auxiliary stochastic process

- For all history h_t, define an auxiliary stochastic process $(\hat{s}_{t+w})$ s.t.
  - Initial value : $\hat{s}_t = s_t$
  - Transitions : $P(\hat{s}_{t+w} = s \mid \hat{s}_{t+w-1}) = \prod_{i=1}^{N} \mu_t^i(\hat{s}^i_{t+w-1} \to s^i)$

  This process is stationary, because it only uses probability of period t.
- We can define

$$\wedge_{t,w} = \sum_{k \in S^i} \mu_T(k \to j) P((k, \hat{s}^{-i})_{t+w} \mid h_T) - P((j, \hat{s}^{-i})_{t+w} \mid h_T)$$

- Now, we want to show that $\wedge$ is close to and that $\wedge$ remains small.

Regret Matching Theorem : Proof

Step 4/8 : and ^ are close

- By definition, we have :

$$(t + v)[D_{t+v}(j;k) \quad D_t(j;k)] = \sum_{w=1} A_{t+w}(j;k) \quad vD_t(j;k)$$

- Using that $jA_{t+w}(j;k)j \quad 2M$ and $jD_{t+w}(j;k)j \quad 2M$:

$$R_{t+v}(j;k) \quad R_t(j;k) = O(\frac{v}{t}) \tag{3}$$

- As R_{t+v} are used for s_{t+v} and R_t for $\hat{s}_{t+v}$, we go from this 1-transition inequality to a w-transitions inequality :

$$\mathrm{P}(s_{t+w} = sjh_t) \quad \mathrm{P}(\hat{s}_{t+w} = sjh_t) = O(\frac{w^2}{t})$$

- Finally, we obtain that and ^ are close :

$$_{t;w}(j;s \quad ^i) \quad \hat{}_{t;w}(j;s \quad ^i) = O(\frac{w^2}{t}) \tag{4}$$

Regret Matching Theorem : Proof

Step 5/8 : Dominate ^

- Denoting $\;_t$ the transition matrix such that $\;_t(i;j) = \;^i_t(i \;\; ! \;\; j)$, we have that

$$P(\hat{s}^i_{t+w} = j|h_t) = \;^w_t(i;j)$$

- By independence, we also have :

$$P(\hat{s}_{t+w} = (j;s^{\;\;-i})|h_t) = P(\hat{s}^{\;\;-i}_{t+w} = s^{\;\;-i}|h_t)P(\hat{s}^i_{t+w} = j|h_t)$$
$$= P(\hat{s}^{\;\;-i}_{t+w} = s^{\;\;-i}|h_t) \;^w_t(i;j)$$

- We inject that into the definition of ^ and we obtain :

$$\wedge_{t;w} = P(\hat{s}^{\;\;-i}_{t+w} = s^{\;\;-i}|h_t)[\;^{w+1}_t \quad \;^w_t](k;j)$$

Regret Matching Theorem : Proof

Step 6/8 : Dominate

Lemma

d0§ @0 9 m m §nH\9 §_H h9§}_² °_§\ 9ee nT _§ L_9Xnj9e 0j§}_0 > 0u ¦ \0j $_{w+1}$ $_{w}$ $= O(\frac{1}{\sqrt{w}})$u

- In the definition of $_t$, we chose a constant c such that $_t(j \; ! \; j) > 0$ in any cases.
- We can then apply the lemma, and we obtain that $\wedge$ tends to 0 with w :

$$\wedge_{t;w} = O(\frac{1}{\sqrt{w}}) \qquad (5)$$

- Using equations (4) and (5), we obtain that

$$_{t;w} = O(\frac{w^2}{t} + \frac{1}{\sqrt{w}}) \qquad (6)$$

Regret Matching Theorem : Proof

Step 7/8 : Dominating the middle term of (1)

- Equations (6) and (2) give us that

$$R_t \quad \mathrm{E}[A_{t+w} j h_t] = O(\frac{w^2}{t} + \frac{1}{p\overline{w}})$$

- It follow from (1) that

$$\mathrm{E}[(t+v)^2 \quad _{t+v} j h_t] = t^2 \quad _T + O(v^3 + t^{p}\overline{v})$$

- $8n; t_n = bn^{\frac{5}{3}}c$, and $v = t_{n+1} \quad t_n = O(n^{\frac{2}{3}})$:

$$\mathrm{E}[t_{n+1}^2 \quad _{t_{n+1}} j h_{t_n}] \quad t_n^2 \quad _{t_n} + O(n^2)$$

Regret Matching Theorem : Proof

Step 8/8 : Show the convergence

- We use ~§}njX d9° nT d9}X0 i«h@0} Tn} K0q0jL0j§ |9jLnh -9}_9@e0 ¦ \0n}0h and obtain $\lim_{n \to \infty} \; _{t_n} = 0$.
- $_{t_n} = \sum_{k \neq j} [R_{t_n}(j;k)]^2$, so $\forall k; j; \lim_{n \to \infty} R_{t_n}(j;k) = 0$.
- For $t_n \quad t \quad t_{n+1}$, we have from (3) that

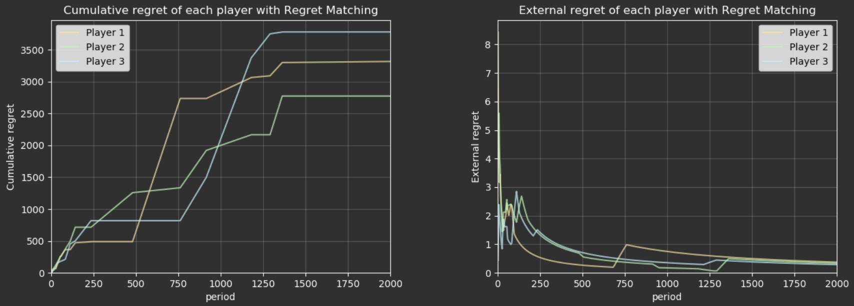
$$R_t(j;k) \quad R_{t_n}(j;k) = O\left(\frac{n^{2/3}}{n^{5/3}}\right) = O\left(\frac{1}{n}\right)$$

- We obtain what we wanted : All internal regrets vanish at the limit

$$\forall j; k; R_t(j;k) \to 0 \qquad (7)$$

- The convergence to the set of correlated equilibrium follows.

Regret Matching : Experiment



S_X«}0 – Evolution of cumulative regret (left) and internal regret (right) with Regret matching

$$\text{G«h«e9§\_®O |OX}O§}(i;T) = \sum_{t=1}^{T} (\max_{k \in S^i} u^i(k; s_t^{-i}) - u^i(s_t^i; s_t^{-i})) \quad (8)$$

Real-life behaviours

- No regret) Play the same action : Ȳi 0®0} H\9jX0 9 °_jj0} §09hȲ
- Some regret) Probability to switch : Ȳ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% hn}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ȳ~§9§« x«n @_9 Ȳ.

Real-life behaviours

- No regret) Play the same action : Ȳ i 0®0} H\9jX0 9 °_jj0} §09hȲ
- Some regret) Probability to switch : Ȳ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% hn}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ȳ~§9§« x«n @_9 Ȳ.

Real-life behaviours

- No regret) Play the same action : Ȳ i 0®0} H\9jX0 9 °_jj0} §09hȲ
- Some regret) Probability to switch : Ȳ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% hn}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ȳ~§9§« x«n @_9 Ȳ.

Real-life behaviours

- No regret) Play the same action : Ÿi0®0}H\9jX0 9 °_jj0} §09hŸ
- Some regret) Probability to switch : Ÿ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% hn}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ÿ~§9§« x«n @_9 Ÿ.

Real-life behaviours

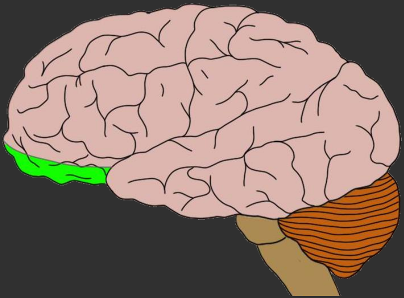
- No regret) Play the same action : Ȳ i 0®0} H\9jX0 9 °_jj0} §09 h Ȳ
- Some regret) Probability to switch : Ȳ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% h n}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ȳ~§9§« x«n @_9 Ȳ.

Real-life behaviours

- No regret) Play the same action : Ȳ i 0®0} H\9jX0 9 °_j j0} §09 h Ȳ
- Some regret) Probability to switch : Ȳ^T ´n« _j®0 §0L _j A _j §09L nT BJ ´n« °_ee \9®0 X9_j0L 20% h n}0 @´ jn°u ~n °_§H\ §n A jn° R
- People tend to have too much inertia) probability to not switch > 0 : Ȳ~§9§« x«n @_9 Ȳ.

In our brain

Regret influences choices in some aspect and is experienced in the orbitofrontal cortex (Camille et al. 2004) [2, 7]



Generalized Regret Matching

Choose a function f verifying the two conditions below and use the following mixed action.

1. f is Lipchitz continous.
2. Sign preserving property : $f(x) > 0$ for $x > 0$ and $f(0) = 0$

$$\sigma_{T+1}(k) = \begin{cases} f(R_T(k)) & \text{if } k \neq j \\ 1 - \sum_{k \neq j} f(R_T(k)) & \text{if } k = j \end{cases}$$

We can also have different $f_{k,j}$ for each $k \neq j$ or allow f to depend on the whole vector of regrets.

Generalized Regret Matching

Choose a function f verifying the two conditions below and use the following mixed action.

1. f is Lipchitz continous.
2. Sign preserving property : $f(x) > 0$ for $x > 0$ and $f(0) = 0$

$$\sigma_{T+1}(k) = \begin{cases} f(R_T(k)) & \text{if } k \neq j \\ 1 - \sum_{k \neq j} f(R_T(k)) & \text{if } k = j \end{cases}$$

We can also have different $f_{k,j}$ for each $k \neq j$ or allow f to depend on the whole vector of regrets.

Generalized Regret Matching Theorem

Theorem (Hart and Mas-Colell, 2001 [1, 3])

d0§ 09H\ qe9´0} qe9´ 9 X0j0}9e_¶0L }0X}0§ h9§H_jX §}9§0X´u ¦ \0j §\0 an_j§ L_ §}_@«§_nj nT qe9´ Hnj®0}X0 §n §\0 0§ nT Hn}}0e9§0L 0x«_e_@}_9 nT §\0 §9X0 X9h0u

:2M2` HBx2/ _2;`2i J i+?BM; , 1tT2`

6B;m`21pQHmiBQM Q7 +mKmH iBp2 `2;`2ib UH27iV _2;`2i K i+?BM;

q2 mb2 i?2 7QHHQrBM;p2+iQ` Q7 7m

$$f_k : (R(1), \ldots, R(n)) \rightarrow C_0 \frac{R(k)}{\sum_i R(i)}$$

Pi?2` `2;`2i@# b2/ bi` i2;B2b

- lM+QM/BiBQM H `2;`2i K i+?BM; q2 iKF V2+i;B2M; : $\frac{1}{T}\sum_{t=1}^{T} u^i(k; s_t^{\,i})$ X
  h?2Q`2K U> `i M/ J b@ 8)Q,/R2(k)H-0ky yy` (HH T
- S`Qtv `2;`2i K il+b2?B M;2M TH v2`b /QMöi FMQ TH vBM; ; K2 M/ QMHv ? p2 ++2bb iQ i?2
$$\hat{R}_{T+1}(k) := \left[\frac{1}{n_k}\sum_{s_t^i = k} u^i(s_t) - \frac{1}{n_j}\sum_{s_t^i = j} u^i(s_t)\right]^+$$
  h?2Q`2K U> `i M/ J b@ 9)Q,H+2QHMl p 2k`y,y2RM(+2 iQ
  TT`QtBK i2 2[mBHB#`B X
- *QMiBMmQmb HHB K2 `,2bmHib + ``v Qp2` QM i
  7` K2rQ`F U> `i M/ J b@e)Q X 2HH- kyyj (

Pi?2` `2;`2i@# b2/ bi` i2;B2b

- lM+QM/BiBQM H `2;`2i K i+?BM; q2 iF2 (k) = $\frac{1}{T}\sum_{t=1}^{T} u^i(k; s_t^{i}) X$
  h?2Q`2K U> `i M/ J b@ 8)V, R_T(k) H - 0 k y y y (HH T
- S`Qtv `2;`2i K i+?BM; lb2 r?2M TH v2`b /QMöi FMQ
  TH vBM; ; K2 M/ QMHv ? p2 ++2bb iQ i?2 B

$$\hat{R}_{T+1}(k) := \left[\frac{1}{n_k} \sum_{s_t^i = k} u^i(s_t) \quad \frac{1}{n_j} \sum_{s_t^i = j} u^i(s_t) \right]^+$$

h?2Q`2K U> `i M/ J b@ 9)V, H+2QHHp 2k`y;y2RM + 2 iQ
TT`QtBK i2 2[mBHB#`B X

- *QMiBMmQmb HHBK2 ;`2bmHib + ``v Qp2` QM i
  7` K2rQ`F U> `i M/ J b@e)V X 2HH- kyyj (

Pi?2` `2;`2i@# b2/ bi` i2;B2b

- lM+QM/BiBQM H `2;`2i KF2i+?BM; q2iK $V_T(k) = \frac{1}{T}\sum_{t=1}^{T} u^i(k; s_t^i)$ X
  h?2Q`2K U> `i M/ J b@8)Q\R2(kH) H-0k7yQy` (HH T
- S`Qtv `2;`2i K i+?BM;r?2,M TH v2`b /QMöi FMQ
  TH vBM; ; K2 M/ QMHv ? p2 ++2bb iQ i?2B

$$\hat{R}_{T+1}(k) := \left[\frac{1}{n_k} \sum_{s_t^i = k} u^i(s_t) - \frac{1}{n_j} \sum_{s_t^i = j} u^i(s_t) \right]^+$$

h?2Q`2K U> `i M/ J b@9)Q\H+2QHMHp 2k`;y2RM(+2 iQ
TT`QtBK i2 2[mBHB#`B X

- *QMiBMmQmb H HBK2 ,2bmHib + ``v Qp2` QM i
  7` K2rQ`F U> `i M/ J b@e)Q\K2HH- kyyj (

amKK `v

R AKTQ`i Mi MQiBQMb
- / TiBp2 >2m`BbiB+b
- LQi iBQMb M/ *QMi2ti
- .Bz2`2Mi FBM/ Q7 2[mBHB#`BmK

k _2;`2i J i+?BM;
- *H bbB+ _2;`2i J i+?BM;
- "2? pBQ` H bT2+ib
- :2M2` HBx2/ _2;`2i J i+?BM;

j / TiBp2 ?2m`BbiB+b M/ L b? 2[mBHB#`BmK
- AMi`Q/m+iBQM
- .vM KB+ bvbi2Kb
- lM+QmTH2/ .vM KB+b h?2Q`2K

AMi`Q/m+iBQM

) Ab i?2`2 / TiBp2 ?2m`BbiB+b i? i L b ?H2 [/miBQHiB?

k@T2`bQM x2`Q@bmK ; K2b
k@T2`bQM TQi2MiB H ; K2b
/QKBM M+2@bQHp #H2 ; K2b
bmT2`KQ/mH ` ; K2b

! Cmbi }m+b 2 iBQmb TH v
U> `i M/ J b@*QH2HH- kyyjV

) q? i #Qmi ;2M2` H ; K2b\

.vM KB+ bvbi2Kb /2}MBiBQM

.2}MBiBQM U.vM KB+ bvbi2KV

/vM KB+ bvbi2K BM+QMiBMmQmb iBK2 ? b i?2 ;2M2

$$\underline{x}(t) = F(x(t);)$$ UN

r?2`2 x2Bb + HHb2i/ i?2 p `B #H2

q2 bb mKK2+QmTH2/ M2ib2 +? TH v2`öb bi` i2;v QrM T vQz ,

$$8i; \underline{x}^i(t) = F^i(x(t); u^i)$$

$$r Bi? x = (x^1; :::; x^N) \quad MF = (F^1; :::; F^N)$$

lM+QmTH2/ .vM KB+b h?2Q`2K

- q2 `2 QMHv bim/vBM; mMKBbmrBiL? b? 2[mBHB#`BmK Q7 x̄( )
- /vM KB+ BbbbL/bi?Q@#+2QMQM8`B87MiU - i?2 mMB L b? 2[mBHB#`BmK Bb
  - R `2bi@TQBMi Q7 i?2 F(x̄( );)=0 B+- BX2X
  - k bi #H2 TQBMi 7Q` i?2 m/vMx(k)B+x̄( ) BX2X 2p2`v bQHmiBQM Q7 U

h?2Q`2K U>`i M/ J b@*QH2HH- kyyj (

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m`

*Q`QHH `v

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m`
+QMp2t ?mHH Q7 i?2 b2i Q7 L b? 2[mBHB#`B X

lM+QmTH2/ .vM KB+b h?2Q`2K

- q2 `2 QMHv bim/vBM; mM+QmTH2/ /vM KB+b
  x̄( )
- /vM KB+ Bb b LB/bi?Q@#+2QMQ 2M`;B?Mi U- i?2 mMB
  L b? 2[mBHB#`BmK Bb
  - R `2bi@TQBMi Q7 i?2 F(x̄( )) ≠ 0 +- BX2X
  - k bi #H2 TQBMi 7Q` i?2 mMx(k) = x̄( ) BX2X 2p2`v
    bQHmiBQM Q7 U

h?2Q`2K U> `i M/ J b@e) QH2HH- kyyj (

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m `

*Q`QHH `v

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m `
+QMp2t ?mHH Q7 i?2 b2i Q7 L b? 2[mBHB#`B X

IM+QmTH2/ .vM KB+b h?2Q`2K

- q2 `2 QMHv bim/vBM; mMKBKmK BL? b? 2-[mMHQBi x̄( )
- /vM KB+ Bb b LB/bi?Q@#+2QMQ2M`;387M2iU - i?2 mMB L b? 2[mBHB#`BmK Bb
  - R `2bi@TQBMi Q7 i?2 F(x̄( ))=0 KB+- BX2X
  - k bi #H2 TQBMi 7Q` i?2 m/vMx(K)B+x̄( )BX7Q2X 2p2`v bQHmiBQM Q7 U

h?2Q`2K U> `i M/ J b@e)QH2HH- kyyj (

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m `

*Q`QHH `v

h?2`2 2tBbi MQ mM+QmTH2/ /vM KB+b i? i ;m `
+QMp2t ?mHH Q7 i?2 b2i Q7 L b? 2[mBHB#`B X

amKK `v

- h?2`2 `2 bBKTH2 / TiBp2 ?2m`BbiB+b i? i H 2[mBHB#`B;`2i J i+?BM;V
- h?2`2 Bb H `;2 +H bb Q7 / TiBp2 ?2m`BbiB +Q``2H i2/ 2[mBHB#`BBx2/ _2;`2i J i+?BM;V
- h?2`2 Bb MQ / TiBp2 ?2m`BbiB+b i? i Hr v 2[mBHB#`B - Q` BM`+QMbBbi2Mi
-) AM b?Q`i- / TiBp2 ?2m`BbiB+b b22K iQ #2 i #2?pBQ` 2H iBQMb +?2bX

amKK `v

- R h?2`2 `2 bBKTH2 / TiBp2 ?2m`BbiB+b i? i H
  2[mBHB#`Bm2;`2i J i+?BM;V
- k h?2`2 Bb H `;2 +H bb Q7 / TiBp2 ?2m`BbiB
  +Q``2H i2/ 2[mBHB#`B x2/ _2;`2i J i+?BM;V
- j h?2`2 Bb MQ / TiBp2 ?2m`BbiB+b i? i Hr v
  2[mBHB#`B - Q` Bb MQM@+QmTH2/ KB+b h?

) AM b?Q`i- / TiBp2 ?2m`BbiB+b b22K iQ #2 i
#2? pBQ`/ 2H iBQM Q +?2bX

amKK `v

- h?2`2 `2 bBKTH2 / TiBp2 ?2m`BbiB+b i? i H
  2[mBHU_#2`B;`2i J i+?BM;V
- h?2`2 Bb H `;2 +H bb Q7 / TiBp2 ?2m`BbiB
  +Q``2H i2/ 2[mBHU_#2`B;`2i J i+?BM;V
- h?2`2 Bb MQ / TiBp2 ?2m`BbiB+b i? i Hr vb
  2[mBHB#`B -Q` BUb M+QQm2tH?2/HvHMXKB+b h?

) AM b?Q`i- / TiBp2 ?2m`BbiB+b b22K iQ #2 i
#2? pBQ`/ 2HH iBQTM`Q +?2bX

amKK`v

- R h?2`2 `2 bBKTH2 / TiBp2 ?2m`BbiB+b i? i H
  2[mBHU_#2`;B`2i J i+?BM;V
- k h?2`2 Bb H `;2 +H bb Q7 / TiBp2 ?2m`BbiB
  +Q``2H i2/ 2[mBHB#`BB x2/ _2;`2i J i+?BM;V
- j h?2`2 Bb MQ / TiBp2 ?2m`BbiB+b i? i Hr vb
  2[mBHB#`B -Q` BUd M+QMp2`;2M+2 iQ h?

) AM b?Q`i- / TiBp2 ?2m`BbiB+b b22K iQ #2 i

#2? pBQ`/ 2HH iBQM`Q +?2bX

. B`2+iBQM Q7 `2b2 `+? bm;;2bi2/

- .Q HH +Q``2H i2/ 2[mBHB#`B `2 Q#i BM2/ + M r2 /2}M2 bK HH2` \ bm#b2i
- q2 FMQr ?Qr i?2b2 bi` i2;B2b #2? p2 BM i?2 #2? pBQM; i?2 r v
- q2 + M mb2 Hi2`M iBp2 `2b2 `+? i2+?MB[m2b 7`Q7 BMbi M iBK2@ p2` ;BM; Q` /Bb+QmMiBM;X
- / TiBp2 ?2m`BbiB+b BMbT`#2/ +QKm+? /Q i?2 `2 H #2? pBQ`b\

.B`2+iBQM Q7 `2b2 `+? bm;;2bi2/

- .Q HH +Q``2H i2/ 2[mBHB#`B `2 Q#i BM2/ + M r2 /2}M2 bK HH2` \bm#b2i
- q2 FMQr ?Qr i?2b2 bi` i2;B2b #2? p2 BM i?2 #2? pBQM; i?2 r v
- q2 + M mb2 Hi2`M iBp2 `2 Q`BMb 7Q` 7BMbi M iBK2@ p2` ;BM; Q` /Bb+QmMiBM;X
- / TiBp2 ?2m`BbiB+b bm+? b Ti `# i2;B2b Kmi+? /Q i? `2 H #2? pBQ`b \

Direction of research suggested

- Do all correlated equilibria are obtained from adaptive heuristics, or can we define a smaller subset ?
- We know how these strategies behave in the limit, but how do they behave along the way ?
- We can use alternative notions of regret, using for instance time-averaging or discounting.
- Adaptive heuristics must be tested in practice : How much do they fit real behaviors ?

Direction of research suggested

- Do all correlated equilibria are obtained from adaptive heuristics, or can we define a smaller subset ?
- We know how these strategies behave in the limit, but how do they behave along the way ?
- We can use alternative notions of regret, using for instance time-averaging or discounting.
- Adaptive heuristics must be tested in practice : How much do they fit real behaviors ?

References I

G9\j, A. General procedures leading to correlated equilibria. ^j§0}j9§_nj9e `n«}j9e nT W9h0 ¦ \0n}´ ¨¨J 21-40 (2004).

G9h_eeO, N. 0§ 9eu The involvement of the orbitofrontal cortex in the experience of regret. ~H_0jH0 ¨•VJ 1167-1170 (2004).

[9}§, S. & g9]GneOee, A. A general class of adaptive strategies.

[9}§, S. & g9]GneOee, A. in NHnjnh_H N 9´ 181-200 (Springer, 2001).

[9}§, S. & g9]GneOee, A. A Simple Adaptive Procedure Leading to Correlated Equilibrium. NHnjnh0§}_H9 £PJ 1127-1150. https://ideas.repec.org/a/ecm/emetrp/v68y2000i5p1127-1150.html (sept. 2000).

[9}§, S. & g9]GneOee, A. Uncoupled dynamics do not lead to Nash equilibrium. 8h0}_H9j NHnjnh_H |0®_0° k¨J 1830-1836 (2003).

References II

¯9ee_ , J. D. Orbitofrontal cortex and its contribution to decision-making. *8jj«u |0®u i0«}n H_u* ¨•J 31-56 (2007).

¦\9jc Tn} ´n«} 9§§Oj§_nj R

github.com/TheoDlmz/AdaptativeHeuristics