Learning from Errors*

Juan Enrique Martínez-Legaz†
CODE and Departament d'Economia i d'Història Econòmica
Universitat Autònoma de Barcelona

Antoine Soubeyran
GREQAM
Université de la Méditerranée

Adress for manuscript correspondence:

Juan Enrique Martínez-Legaz
Departament d'Economia i d'Història Econòmica
Universitat Autònoma de Barcelona
08193 Bellaterra
Spain

E-mail: JuanEnrique.Martinez@uab.es
Phone: (34) 93 581 13 66
Fax: (34) 93 581 20 12

January 21, 2003

*We are grateful to Bruno Soubeyran for stimulating and helpful comments; in particular, for bringing reference [7] to our attention.

†Partial support by the Ministerio de Ciencia y Tecnología, project BEC 2002-00642, and by the Comissionat per Universitats i Recerca de la Generalitat de Catalunya, Grant SGR2001-00162, is gratefully acknowledged. This work has been carried out during several stays I have made at GREQAM; I am grateful to the LEA "Quantitative Economics and Theory of Choice" and to the Université de la Méditerranée for the support received.

Abstract

We present a model of learning in which agents learn from errors. If an action turns out to be an error, the agent rejects not only that action but also neighboring actions. We find that, keepng memory of his errors, under mild assumptions an acceptable solution is asymptotically reached. Moreover, one can take advantage of big errors for a faster learning.

Key words: Learning, errors, fixed points, normal form games, best reply functions.

JEL Classification Numbers: D83, C72.

1 Introduction: A Generalized Tatonnement (or Fictitious Play) Process

In his excellent survey on Learning; Sobel [20] raises the question: “What is a sensible starting point for a learning model?” First of all, let us adopt the following model of ”Learning”: A learning process takes as given that active agents, while doing actions, cannot fully escape to make errors, even if they know that they will make errors, and, period after period, they try to correct their errors. The central points are: first, to define what is an error, and then to be more precise on how agents correct them.

To define errors, models necessarily must specify what agents initially know and how they build on this knowledge. We must explain why agents make bad decisions. Some of them are, for example, because they are inexperienced, ignorant, or face difficult problems.

To model the process of elimination of errors, models must specify what agents learn at each period, how experience accumulates over time and how experience is used to correct errors and make better decisions. The major problem is to know when learning leads to good, acceptable results, in a realistic amount of time.

Our answer comes from smart businessmen with whom we have the opportunity to discuss on their working life experience, asking them why they have succeeded so well. One major question was: Do they optimize?.Most of them say that they experience a succession of successes and failures but that, in their own case, successes where more numerous than failures. None of them thinks that failures where unnecessary. On the contrary, they tell us that failures help them very much not to make a lot of other mistakes! They learn by essays and errors, because they do not fail to correct their errors! One of the most famous French businessmen (Mr. Bellon, Vice President of the “Patronat Français” at the time, 1998), and one of his colleagues from Marseille said this in a very nice form, at the occasion of a conference that one of us organized in Marseille, where reputed businessmen and smart researchers got together to discuss their respective experiences about the roots of success. Mr. Bellon, speaking to the audience, declared: “At the beginning of my working life, I experienced several bosses who where very bad....”. The audience laughed a lot. But Mr. Bellon reacted: “Please, it is not a joke, it was my chance, because, very soon in my working life, I learned what I must not do!”. The problem for him was not to know what to do exactly (this is too difficult in a short time) but to know, as quickly as possible, what kind of big errors to eliminate. We name this the “Solutions Cannot Be There” principle of learning, because it tells us that what is important is to know, as soon as possible, what kind of solutions one must avoid to use for solving some problems (errors that should not be repeated). This is a cumulative process of learning from past errors, based on the well known saying: “Errare humanum est, perseverare diabolicum”. The expression “Solutions Cannot Be There” represents a geometric description of the process of learning. If an agent (or a group of players) tries a solution x in a given space and if, after inspection, it appears that this point is not acceptable (is not a good solution), he (or they) will learn that all points around x (or in a more general "rejected set" linked to x) are not acceptable. Then the "Solutions Cannot Be There" principle defines a process of elimination of non acceptable points.

Our paper poses the question: Does a process of essays and errors, with feedbacks to experience, help to reach an acceptable point (a solution) in a reasonable amount of time? This seems to be the favorite approach that the French ANVAR Institution (2000) promotes to help the emergence and the final success of start-ups.

The problem is to know how much we must learn at each step to get enough information for reaching an acceptable point. But what do we learn? Most often, by doing some actions, we learn that these actions and related ones are not satisfactory (they are mistakes).

Agents with no information on their environment only satisfy. All they can observe is their own and all other agents' actions, immediately after these actions have been done. We model satisfying by defining an acceptable set A; for example, the set of fixed points of the best reply mapping of a normal form game.

In this paper we analyze Learning as a process of essays and errors of one agent or several interacting agents, who do not know a priori what is the best for the others (they do not know the objective functions of other players). Then they experiment by essays and errors to gain information. At each step they learn that what they have done just before (one step before) was not the best answer to their environment (meaning, for example, what others players have done just before). Having done something (one action), given the daily environment, reveals that this action and some other neighboring ones are mistakes. Learning means that, step by step, the agent takes into account this new information (feedback from experiment) to never repeat these errors, which are deleted from the feasible action set. Then at each step the agent learns either that he has done some acceptable action (satisfying) or else how to do better, but, more often, what errors he should never repeat (unsatisfying actions). He deletes dominated actions (mistakes) from the feasible action set. At each period he learns that the best he wants to achieve (his goal, or an acceptable outcome) cannot be in some set, now and forever. When he has done a feasible action, he learns either that it is acceptable (even, less often, optimal) or, if not, that a neighborhood of it is also unacceptable (resp., not optimal). Then, at each period, given the action he has chosen either he learns that this action is acceptable (or optimal), or else he can define a "not to be there" set of points (a rejected set of actions, including the one just chosen), which can be rejected immediately at this step as not acceptable solutions, now and in the future. This is the way the agent acquires information on the successive errors, which he should not repeat, each time he makes a wrong choice. Step by step, he narrows the feasible action set until he does something acceptable or optimal.

There are two main messages in this paper. The first one is that learning step by step some actions that have to be avoided, even if there is no clear vision of the acceptable solutions of a problem, may be fruitful for achieving one of them. The second one is that big errors may be more useful than smaller ones, because they allow us to learn more in the case of continuous learning processes; in this case, we can exhibit a linking mechanism between the size of errors and the amount of learning.

2 A Learning Process by Local Elimination Around Unsuccessful Essays

In this section we build a model for the first message this paper conveys. We consider a very simple framework for decision processes and propose an iterative algorithm that produces a sequence of actions, after making each of which one learns, in case it is not acceptable, that a whole set of other neighboring actions is unacceptable, too. An agent who has the capability of learning will keep memory of these rejectable actions throughout the whole process; under mild assumptions, an acceptable solution will asymptotically be reached.

We consider an agent facing a simple decision process $((X,d),A)$, which consists of a nonempty metric space of possible actions (X,d) and a closed set of acceptable solutions $A \subset X$, which the agent does not know until he reaches a point in it.

To each unacceptable action $x \in X\backslash A$ that the agent takes, let us associate an open set $N(x)$. We interpret this set as consisting of those points that the agent is certain not to be acceptable after realizing that action x was not a good decision. If $x \in A$ we set $N(x) = \emptyset$. Once the agent reaches a point x, two cases can happen. The first case occurs when he learns that x is acceptable: $x \in A$.and $N(x) = \emptyset$. We will impose to the mapping $N : X \rightarrow 2^X$ the condition that $N(x) = \emptyset$ implies $x \in A$ (otherwise we would not be able to reject a point x, that is, to decide that $x \notin A$, just by knowing that $N(x) = \emptyset$). The second case occurs when the agent learns that point x is not acceptable, one then has $x \in N(x)$, so that he also learns that a neighborhood of x is not acceptable. Thus, if he makes a mistake, he learns a lot. In future periods he will never come back to such a rejected subset $N(x)$ for choosing the next action. This is the exact transcription of the sentences "Errare humanum est, perseverare diabolicum", and "try to never repeat the same error". This means that if at period t, $F_t \subset X$ is the subset of "not yet rejected points", at period $t+1$ the subset of not yet rejected points will be $F_{t+1} = F_t \backslash N(x)$. At the beginning of the learning process, to reach an acceptable point, the set of not yet rejected points is $F_0 = X$. Then we have the successive inclusions $F_0 \supset .... \supset F_t \supset F_{t+1} \supset A$, with $x_t \in F_t$ for each $t = 0, 1, 2, ...t$. We see very clearly that this learning process contracts more and more the not rejected set $F_t \supset A$, hoping to reach one point of the acceptable set of points A. But we do not know whether the process does not stop, becomes almost stationary,and remains always at a distance of A greater than some strictly positive constant $d_* > 0$. We will show that this is not the case, under the very mild assumption that whenever the agent reaches a non acceptable point $x \notin A$ then he learns that a ball of a strictly positive radius of neighboring points is also not acceptable. The striking result will be that this radius can be as small as one can imagine. In fact there is no size condition on it, except that the radius depend in a lower semicontinuous way on the point. So, in case of indefinite successive errors, the agent needs to learn very little at each period to succeed at the end. A lot of errors gives a lot of information.

Definition 1 *Let (X,d) be a nonempty metric space and $A \subset X$. A learning mapping of rejection for the decision problem $((X,d),A)$ is any mapping $N : X \to 2^X$ with open images satisfying*

i) $x \in A \Rightarrow N(x) = \emptyset$.

ii) $x \notin A \Rightarrow x \in N(x) \subset X \backslash A$.

We assume that the agent has an available learning mapping of rejection $N : X \to 2^X$ such that, after taking an action $x \in X$, he can easily compute $N(x)$. In this way he learns that either x is acceptable ($N(x) = \emptyset$) or else that there is an open neighborhood $N(x)$ of x that does not contain any acceptable solution. We shall use the function $r_N : X \to \mathbb{R}_+$ given by $r_N(x) = d(x, X\backslash N(x))$ to measure the degree of undesirability of points in X. Notice that $r_N(x) = 0$ if and only if $x \in A$. In fact, $r_N(x)$ is the distance to the closest point to x that is not rejected at x. In the special case when the rejected sets are balls, $r_N(x)$ is greater than or equal to the radius of $N(x)$. We shall assume that the function r_N satisfies the following mild continuity condition:

$$\text{If } \overline{x} \text{ is a limit point of } X \text{ and } \lim_{x \to \overline{x}} r_N(x) = 0 \text{ then } r_N(\overline{x}) = 0. \tag{1}$$

One can easily prove that this condition is equivalent to the existence of a nonnegative lower semicontinuous minorant of r_N that vanishes only on A. Obviously, this condition holds if and only if the lower semicontinuous hull of r_N (that is, the largest lower semicontinuous minorant of r) vanishes only on A; in particular, it holds if r_N is lower semicontinuous. Thus, by [2, Chapter 4, Section 8, Theorem 4], one has

Proposition 2 *Let (X,d) be a nonempty metric space, $A \subset X$ and $N : X \to 2^X$ be a learning mapping of rejection for the decision problem $((X,d),A)$. If X is compact and the mapping* $\begin{array}{c} X \to 2^X \\ x \longmapsto X\backslash N(x) \end{array}$ *is upper semicontinuous then r_N is lower semicontinuous and therefore condition (1) holds.*

Relative to a learning mapping of rejection $N : X \to 2^X$, we define the following local learning process by successive essays and errors:

$(LLP)_N$
Step 0. Set $t = 0$ and $F_t = X$
Step 1. If $F_t = \emptyset$, then $A = \emptyset$. Stop.
Step 2. If $F_t \neq \emptyset$, choose $x_t \in F_t$.

Step 3. If $N(x_t) = \emptyset$, then $x_t \in A$. Stop.
Step 4. If $N(x_t) \neq \emptyset$, define $F_{t+1} = F_t \backslash N(x_t)$.. Increase t by 1 and go to Step 1.

We have the inclusions $F_0 \supset .... \supset F_t \supset F_{t+1} \supset A$, with $x_t \in F_t$ for each $t = 0, 1, 2, ...$ Indeed, the inclusion $F_{t+1} \supset A$ follows from the fact that $N(x_t) \subset X \backslash A$ for all $x \in X$. This inclusion guarantees that if $(LLP)_N$ stops at Step 1 then $A = \emptyset$. Since N is also assumed to satisfy $N(x) = \emptyset$ only for $x \in A$, we are also certain that $(LLP)_N$ yields an acceptable solution if it stops at Step 3. The case when $(LLP)_N$ does not stop is considered next:

Theorem 3 *Let (X, d) be a nonempty metric space, $A \subset X$ and $N : X \to 2^X$ be a learning mapping of rejection for the decision problem $((X, d), A)$ such that r_N satisfies (1).*

If $(LLP)_N$ does not stop then every limit point of the sequence $\{x_t\}$ generated by $(LLP)_N$ belongs to A; if, in addition, X is compact then $r_N(x_t) \to 0$.

Proof. Suppose that $\overline{x}$ is a limit point of $\{x_t\}$, so that $\overline{x} = \lim x_{t_k}$ for some subsequence $\{x_{t_k}\}$. Notice that $\overline{x} \neq x_{t_k}$ for each k (since $x_{t_l} \in F_{t_l} \subset F_{t_k} = X \backslash N(x_{t_k})$ for $l > k$), so that $\overline{x}$ is a limit point of X, too. For every k one has $x_{t_{k+1}} \in F_{t_{k+1}} \subset F_{t_k+1} = F_{t_k} \backslash N(x_{t_k}) \subset X \backslash N(x_{t_k})$, whence $d(x_{t_k}, x_{t_{k+1}}) \geq d(x_{t_k}, X \backslash N(x_{t_k})) = r_N(x_{t_k}) \geq 0$. Therefore, given that $d(x_{t_k}, x_{t_{k+1}})$ tends to 0 (as $\{x_{t_k}\}$ converges), it follows that $r_N(x_{t_k})$ tends to 0, too, which, by (1), implies that $r_N(\overline{x}) = 0$, i.e. $\overline{x} \in A$.

If X is compact then r_N, being nonnegative and majorized by the continuous function $x \longmapsto d(x, A)$, is bounded and hence, to prove that $r_N(x_t)$ tends to 0, it suffices to prove that any convergent subsequence $r_N(x_{t_k})$ tends to 0. But $\{x_{t_k}\}$ has a convergent subsequence $\left\{x_{t_{k_l}}\right\}$, which tends toward some $\overline{x} \in A$. We thus have $0 \leq \lim_{k \longrightarrow \infty} r_N(x_{t_k}) = \lim_{k \longrightarrow \infty} r_N(x_{t_{k_l}}) \leq \lim_{k \longrightarrow \infty} d(x_{t_{k_l}}, A) = d(\overline{x}, A) = 0$. ■

According to the preceding theorem, three cases can occur. One of them is that at some iteration t we learn that the feasible set F_t is empty, which implies that the acceptable set A is empty, too. This reveals that our aspirations are unrealistic, and so we should revise them to make them less demanding. The most fortunate possibility corresponds to the case when at some iteration we obtain an acceptable solution. If none of these two cases occur, the process does not stop and then we can always hope to improve our current solution. Indeed, in this case, in spite of the fact that convergence is not guaranteed even if the space is compact, we know that every convergent subsequence has its limit point in the acceptable set; in other words, we are not following a wrong path. The need to consider a subsequence to get convergence means that one might have to disregard infinitely many actions (unsuccessful trials) taken during the process. Compactness of the space of actions is a sufficient condition for the existence of such a convergent subsequence to an acceptable solution; the role of compactness here has the intuitive interpretation that the "size" (or the diameter) of the non rejected set $F_t = X \backslash \bigcup_{k=0}^{t-1} N(x_t)$ decreases at each iteration (notice, for instance, that when X is unbounded all the rejected sets may also be unbounded even if the acceptable set is very small). Surprisingly, convergence does not require large rejected sets $N(x_t)$, as our theorem does not contain any assumption on their size; thus we get convergence of a subsequence (under compactness) even if we reject a very small set at each iteration. This is in the spirit of bounded rationality, as it says that it is enough to get just a little local information at each iteration. Moreover, compactness also ensures that the degree of undesirability tends to zero, so that a sufficiently advanced iterate has a degree of undesirability as little as we are ready to tolerate.

Corollary 4 *Under the assumptions of Theorem 3, if $(LLP)_N$ does not stop and X is nonempty and compact then A is nonempty.*

Proof. We have $F_0 = X \neq \emptyset$. Thus either $(LLP)_N$ stops at some $x_t \in A$ or else it generates a sequence $\{x_t\}$, which, X being compact, must have some limit point $\overline{x} \in X$. According to Theorem 3, $\overline{x} \in A$. ■

Corollary 5 *Under the assumptions of Theorem 3, if $(LLP)_N$ does not stop, X is compact and A is a singleton, $A = \{a\}$, then the sequence $\{x_t\}$ generated by $(LLP)_N$ converges to a.*

Proof. It suffices to observe that, according to Theorem 3, a is the unique limit point of $\{x_t\}$. ■

3 Linking the Amount of Learning to the Size of Errors: Continuous Learning Processes

The aim of this and the following sections is to propose a more precise model for the second message of this paper. We consider the case, frequent in economic models, when the acceptable solutions of a decision process are the fixed points of a continuous mapping from a metric space into itself. In this case we can relate the amount of learning at each possible action to the size of the error made when taking it. We propose a method to reject a neighborhood of a wrong action that gives rise to rejected sets with sizes that depend on the sizes of errors in a nondecreasing way.

Let (X,d) and (X',d') be metric spaces and $f : X \to X'$ be a mapping. We recall that $\eta : X \times (0,+\infty) \to (0,+\infty]$ is called a modulus of continuity for f if for every $\varepsilon > 0$ and $x, y \in X$ such that $d(x,y) < \eta(x,\varepsilon)$ one has $d'(f(x), f(y)) < \varepsilon$. Clearly, the existence of a modulus of continuity for f is equivalent to the continuity of f. Indeed, if f is continuous then one can define a modulus of continuity $\overline{\eta}$ by

$$\overline{\eta}(x,\varepsilon) = \inf\{d(x,y) \ / \ y \in X, \ d'(f(x), f(y)) \geq \varepsilon\};$$

continuity of f at $x \in X$ guarantees that $\overline{\eta}(x,\varepsilon) > 0$ for all $\varepsilon > 0$. One can easily check that $\overline{\eta}$ is nondecreasing in ε and is the (pointwise) largest modulus of continuity of f. In fact, it is not hard to prove that $\eta : X\times(0,+\infty) \to (0,+\infty]$ is a modulus of continuity for f if and only if $\eta \leq \overline{\eta}$. The following result and its surprisingly simple proof, which we reproduce here with a slightly different presentation, can be found in [7]:

Proposition 6 *Let (X,d) and (X',d') be metric spaces and $f : X \to X'$ be a continuous mapping. There is a modulus of continuity $\eta : X\times(0,+\infty) \to (0,+\infty)$ for f that satisfies the following conditions:*

i) *η is Lipschitzian*[1] *with constant 1, the space $X\times(0,+\infty)$ being endowed with the box metric d'' defined by $d''((x,\varepsilon),(x',\varepsilon')) = \max\{d(x,x'),|\varepsilon-\varepsilon'|\}$.*

ii) *η is nondecreasing in its second argument.*

Proof. Define η by

$$\eta(x,\varepsilon) = \inf_{u\in X, v\in X} \max\{d(x,u), d(x,v), \varepsilon - d'(f(u), f(v))\}.$$

Let us first observe that η is well defined, that is, it takes only strictly positive values. Let $(x,\varepsilon)\in X\times(0,+\infty)$ and take $\delta>0$ such that $d'(f(x),f(y)) \leq \frac{\varepsilon}{3}$ for every $y\in X$ satisfying $d(x,y)<\delta$. Then if $u\in X$ and $v\in X$ are such that $d(x,u)<\delta$ and $d(x,v)<\delta$, one has $d'(f(u),f(v)) \leq d'(f(x),f(u)) + d'(f(x),f(v)) < \frac{2\varepsilon}{3}$, whence $\varepsilon - d'(f(u),f(v)) > \frac{\varepsilon}{3}$. *This* proves that, for every $u\in X$ and $v\in X$, $\max\{d(x,u),d(x,v),\varepsilon-d'(f(u),f(v))\} \geq \min\{\delta,\frac{\varepsilon}{3}\}$; therefore $\eta(x,\varepsilon) \geq \min\{\delta,\frac{\varepsilon}{3}\} > 0$.

To prove *i)*, observe that, for any $u,v\in X$, the functions $(x,\varepsilon)\mapsto d(x,u)$, $(x,\varepsilon)\longmapsto d(x,v)$ and $(x,\varepsilon)\mapsto \varepsilon - d'(f(u),f(v))$ are Lipschitzian with constant 1 on $X\times(0,+\infty)$, and recall that the class of Lipschitzian functions with a given constant is closed under pointwise suprema and infima.

Condition *ii)* is obvious. ■

There is no loss of generality in assuming that a modulus of continuity, besides being Lipschitzian with constant 1 and nondecreasing in its second argument, satisfies $\widehat{\eta}(x,\varepsilon)\leq\varepsilon$ for every $x\in X$ and $\varepsilon>0$, since one can define $\widehat{\eta} : X\times(0,+\infty)\to(0,+\infty)$ by $\widehat{\eta}(x,\varepsilon) = \min\{\eta(x,\varepsilon),\varepsilon\}$, $\eta : X\times(0,+\infty)\to(0,+\infty)$ being as in the statement of the preceding proposition. The next theorem will require the modulus of continuity to have this additional property, but will use only continuity in the first argument (rather than the joint Lipschitz condition).

Theorem 7 *Let (X,d) be a metric space, $f : X\to X$ be a continuous mapping, $\varepsilon\geq 0$, and suppose that the set*

$$F_\varepsilon(f) = \{x\in X \ / \ d(f(x),x)\leq\varepsilon\}$$

is nonempty. If $\eta : X\times(0,+\infty)\to(0,+\infty)$ is a modulus of continuity for f that is continuous in its first argument and satisfies $\eta(x,\varepsilon)\leq\varepsilon$ for every $x\in X$

[1] See the beginning of the next section for the definition of the notion of Lipschitzian mapping.

and $\varepsilon > 0$ then

$$d(x, F_\varepsilon(f)) \geq \eta \left(x, \frac{1}{2} (d(x, f(x)) - \varepsilon) \right) \text{ for all } x \in X \tag{2}$$

(here and in the sequel we use the convention $\eta(x,t) = 0$ if $t \leq 0$).

Therefore, if the set A of acceptable solutions consists of the fixed points of f, $A = F_0(f)$, the mapping $N : X \to 2^X$ defined by

$$N(x) = \begin{cases} \emptyset & \text{if } x \in A \\ B(x, r(x)) & \text{if } x \notin A \end{cases},$$

$B(x, r(x))$ denoting the open ball with center x and radius

$$r(x) = \eta \left(x, \frac{1}{2} d(x, f(x)) \right), \tag{3}$$

is a learning mapping of rejection for the decision problem $((X,d), A)$; moreover, the function r_N satisfies condition (1). Hence, if η is nondecreasing in its second argument and $(LLP)_N$ does not stop then every limit point of the sequence $\{x_t\}$ generated by $(LLP)_N$ is a fixed point of f; if, in addition, X is compact then $d(x_t, f(x_t)) \to 0$.

Proof. Let $x \in X \backslash F_\varepsilon(f)$. By definition,

$$d(x, F_\varepsilon(f)) = \inf \{ d(x,y) \ / \ y \in F_\varepsilon(f) \}.$$

Let $y \in X$ be such that $d(x,y) < \eta \left(x, \frac{1}{2} (d(x, f(x)) - \varepsilon) \right)$. We have

$$\begin{aligned} d(x, f(x)) &\leq d(x,y) + d(y, f(y)) + d(f(y), f(x)) \\ &< \eta \left(x, \frac{1}{2} (d(x, f(x)) - \varepsilon) \right) + d(y, f(y)) + \frac{1}{2} (d(x, f(x)) - \varepsilon) \\ &\leq \frac{1}{2} (d(x, f(x)) - \varepsilon) + d(y, f(y)) + \frac{1}{2} (d(x, f(x)) - \varepsilon) \\ &= d(x, f(x)) - \varepsilon + d(y, f(y)), \end{aligned}$$

whence $d(y, f(y)) > \varepsilon$, i.e. $y \notin F_\varepsilon(f)$. This proves (2).

We shall now show that, if η is nondecreasing in its second argument, r_N satisfies (1). Assume that $\overline{x}$ is a limit point of X such that $\lim_{x \to \overline{x}} r_N(x) = 0$, and let $\varepsilon > 0$. One has $\overline{x} = \lim x_k$ for some sequence $\{x_k\} \subset X \backslash \{\overline{x}\}$. If some subsequence of $\{x_k\}$ is contained in A then, as A is closed, $\overline{x} \in A$, that is, $r_N(\overline{x}) = 0$. So we can assume that $x_k \notin A$ for every sufficiently large k. Since $\lim \left(r_N(x_k) - \eta \left(x_k, \frac{1}{2} \varepsilon \right) \right) = -\eta \left(\overline{x}, \frac{1}{2} \varepsilon \right) < 0$, for k large enough one has $\eta \left(x_k, \frac{1}{2} d(x_k, f(x_k)) \right) = r(x_k) \leq r_N(x_k) < \eta \left(x_k, \frac{1}{2} \varepsilon \right)$, whence, as η is nondecreasing in its second argument, $d(x_k, f(x_k)) < \varepsilon$. Therefore, by continuity, $d(\overline{x}, f(\overline{x})) \leq \varepsilon$. Since ε is an arbitrary positive number, it follows that $d(\overline{x}, f(\overline{x})) = 0$, i.e., $\overline{x} \in A$ or, equivalently, $r_N(\overline{x}) = 0$. We have thus proved the claim that r_N satisfies (1). Therefore, by Theorem 3, every limit point of the sequence $\{x_t\}$ generated by $(LLP)_N$ belongs to A.

To finish the proof assume, towards a contradiction, that $d(x_t, f(x_t))$ does not converge to 0, that is, that there is a positive number ε and a subsequence $\{x_{t_k}\}$ of $\{x_t\}$ such that $d(x_{t_k}, f(x_{t_k})) \geq 2\varepsilon$ for every k. Since η is nondecreasing in its second argument and strictly positive, using again Theorem 3 we get $0 < \eta(x_{t_k}, \varepsilon) \leq \eta\left(x_{t_k}, \frac{1}{2}d(x_{t_k}, f(x_{t_k}))\right) = r(x_{t_k}) \leq r_N(x_{t_k}) \to 0$, whence $\eta(x_{t_k}, \varepsilon) \to 0$. But, since X is compact, $\{x_{t_k}\}$ has a further subsequence $\left\{x_{t_{k_l}}\right\}$ that converges to some point $\overline{x} \in A$. Given that η is continuous in its first argument, this yields the contradiction $0 = \lim \eta(x_{t_k}, \varepsilon) = \lim \eta\left(x_{t_{k_l}}, \varepsilon\right) = \eta(\overline{x}, \varepsilon) > 0$. ■

The assumption that the modulus of continuity $\eta(x, \varepsilon)$ is nondecreasing in ε, required in the proof of the convergence part of the preceding theorem, admits the following interpretation. If we regard $d(x, f(x))$ as the error at x (indeed, it measures the extent to which x fails to be a fixed point) then, by (3), the radius of the rejected ball at x depends on the error at x in a nondecreasing way. Put in other words, the bigger the mistake is the more one learns. So we have a clear link between the size of errors and the amount of learning, which illustrates our second message in this paper that big errors are fruitful because one learns a lot.from them. On the other hand our convergence result indicates that, under compactness and continuity, the sequence of errors converges to zero, so that learning is efficient in this case.

Corollary 8 *Let (X, d), f, η and N be as in Thm. 7. If η is nondecreasing in its second argument, $(LLP)_N$ does not stop, X is compact and f has a unique fixed point a then the sequence $\{x_t\}$ generated by $(LLP)_N$ converges to a.*

We want to emphasize that in the case the preceding results deal with, that is, when the acceptable solutions are the fixed points of a continuous mapping f, our algorithm $(LLP)_N$ is different from the typical iteration process $x_{t+1} = f(x_t)$. Notice that, in case of existence, our algorithm always converges to a fixed point of f while the iteration process need not converge. The inefficiency of the iteration process in case of non convergence lies therefore in the fact that at some iteration it visits some point which was known to be non acceptable at some previous iteration. The inefficiency of the iteration process in the search of fixed points and the need to modify it to get convergence is known since long time ago (see, e.g., [14]).

Recall that a mapping $f : X \to X'$ between two metric spaces is uniformly continuous when it has a modulus of continuity $\eta : X \times (0, +\infty) \to (0, +\infty]$ that does not depend on its first argument, that is, $\eta(x, \varepsilon)$ depends only on ε. In this case the function $\omega : (0, +\infty) \to (0, +\infty]$ defined by $\omega(\varepsilon) = \eta(x, \varepsilon)$ is said to be a modulus of uniform continuity for f. The (pointwise) largest modulus of uniform continuity $\overline{\omega}$ of f is given by

$$\overline{\omega}(\varepsilon) = \inf\{d(x, y) \;\; / \;\; d'(f(x), f(y)) \geq \varepsilon\}.$$

Clearly, $\overline{\omega}$ is nondecrasing; moreover, $\omega : (0,+\infty) \to (0,+\infty]$ is a modulus of uniform continuity for f if and only if $\omega \leq \overline{\omega}$. Notice that any modulus of uniform continuity, when regarded as a function of both x and ε, obviously satisfies the continuity assumption required in Theorem 7. The same argument used above shows that there is no loss of generality in assuming that a modulus of continuity ω is nondecreasing and satisfies $\omega(\varepsilon) \leq \varepsilon$ for every ε. Thus, in the case when $f : X \to X$ is uniformly continuous, Theorem 7 and Corollary 8 remain valid and meaningful if they are restated in terms of a modulus of uniform continuity in place of a modulus of continuity.

4 Lipschitzian Learning Processes

The aplicability of the results of the preceding section requires knowing a modulus of continuity of f satisfying suitable properties, as the construction of rejected balls uses this information. Though this might be difficult in general, one often deals with Lipschitzian mappings with a known Lipschitz constant. In this case a simple appropriate modulus of continuity is at hand. We recall that a mapping $f : X \to X'$ between two metric spaces (X, d) and (X', d') is said to be Lipschitzian with constant $C \geq 0$ if $d'(f(x), f(y)) \leq Cd(x,y)$ for every $x, y \in X$. Clearly, this is equivalent to saying that the function $\widehat{\omega} : (0,+\infty) \to (0,+\infty)$ given by $\widehat{\omega}(\varepsilon) = \frac{\varepsilon}{C}$ is a modulus of uniform continuity for f. Thus, in this case we can apply Theorem 7 with the modulus of continuity η defined by $\eta(x,\varepsilon) = \frac{\varepsilon}{\max\{C,1\}}$. By doing so, (2) can be written as

$$d(x, F_\varepsilon(f)) \geq \frac{1}{2\max\{C,1\}} \left(d(x, f(x)) - \varepsilon\right) \text{ for all } x \in X,$$

and we can define a learning mapping of rejection $N : X \to 2^X$ by $N(x) = \begin{cases} \emptyset & \text{if } x \in A \\ B(x, r(x)) & \text{if } x \notin A \end{cases}$, $B(x, r(x))$ denoting the open ball with center x and radius $r(x) = \frac{1}{2\max\{C,1\}} d(x, f(x))$. In the case when $C > 1$, the following theorem gives a sharper lower bound for the distance function to the set of approximate fixed points of f and provides a larger rejection mapping:

Theorem 9 *Let (X, d) be a metric space, $f : X \to X$ be a Lipschitzian mapping with constant C, and suppose that the set*

$$F_\varepsilon(f) = \{x \in X \ / \ d(f(x), x) \leq \varepsilon\}$$

is nonempty. Then

$$d(x, F_\varepsilon(f)) \geq \frac{1}{1+C} \left(d(x, f(x)) - \varepsilon\right) \text{ for all } x \in X. \tag{4}$$

Therefore, if the set A of acceptable solutions consists of the fixed points of f, $A = F_0(f)$, the mapping $N : X \to 2^X$ defined by

$$N(x) = \begin{cases} \emptyset & \text{if } x \in A \\ B(x, r(x)) & \text{if } x \notin A \end{cases},$$

$B(x, r(x))$ denoting the open ball with center x and radius

$$r(x) = \frac{1}{1+C} d(x, f(x)),$$

is a learning mapping of rejection for the decision problem $((X,d), A)$; moreover, the function r_N satisfies condition (1). Hence, if $(LLP)_N$ does not stop then every limit point of the sequence $\{x_t\}$ generated by $(LLP)_N$ is a fixed point of f; if, in addition, X is compact then $d(x_t, f(x_t)) \to 0$.

Proof. Let $x \in X \setminus F_\varepsilon(f)$. By definition,

$$d(x, F_\varepsilon(f)) = \inf\{d(x,y) \;/\; y \in F_\varepsilon(f)\}.$$

For every $y \in F_\varepsilon(f)$ we have

$$\begin{aligned} d(x,f(x)) &\leq d(x,y) + d(y,f(y)) + d(f(y),f(x)) \leq d(x,y) + \varepsilon + Cd(y,x) \\ &= \varepsilon + d(1+C)\, d(x,y), \end{aligned}$$

whence $d(x,y) \geq \frac{1}{1+C}\left(d(x,f(x)) - \varepsilon\right)$; therefore (4) holds.

To prove that r_N satisfies condition (1), let $\overline{x}$ be a limit point of X such that $\lim_{x \to \overline{x}} r_N(x) = 0$. As in the proof of Thm 7, we can assume that $\overline{x} = \lim x_k$ for some sequence $\{x_k\} \subset X \setminus A$. As $0 \leq d(x_k, f(x_k)) = (1+C)\, r(x_k) \leq (1+C)\, r_N(x_k) \to 0$, one has $d(\overline{x}, f(\overline{x})) = \lim_{k\to\infty} d(x_k, f(x_k)) = 0$, so that $\overline{x} \in A$, that is, $r_N(\overline{x}) = 0$. Using Thm. 3 and the inequalities $0 \leq d(x_t, f(x_t)) \leq (1+C)\, r_N(x_t)$, the last assertion of the statement follows. ■

Corollary 10 *Let (X,d) and f be as in Thm. 9. If $(LLP)_N$ does not stop, X is compact and f has a unique fixed point a then the sequence $\{x_t\}$ generated by $(LLP)_N$ converges to a.*

5 Contracting Learning Processes

Contractions are a very important class of Lipschitzian mappings. We recall that a mapping $f : X \to X$ from a metric space (X,d) into itself is said to be a contraction if it is Lipschitzian with constant $0 \leq \alpha < 1$. The Cournot model "with usual assumptions" has a contracting best reply function, as Long and Soubeyran [13] have shown recently. This demonstrates the importance of contractions for dynamic processes of convergence towards equilibrium.

The next theorem shows that, when the acceptable set consists of the fixed points of a contraction, we can make a more powerful rejection process than in the general Lipschitzian case, since then we are able to reject not only a ball centered at x but also the complement of another larger ball, so that the feasible set at each iteration is obtained from that at the previous one by making the intersection with an annulus.

Theorem 11 *Let (X,d) be a complete metric space, $f : X \to X$ be a contraction,with constant α, and $\overline{x}$ be the fixed point of f. For each $x \in X$, let*

$$C(x) = \left\{ y \in X \ / \ \frac{1-\alpha}{1+\alpha} d(x, f(x)) \leq d(x,y) \leq \frac{1}{1-\alpha} d(x, f(x)) \right\}.$$

Then, for every $x \in X$, one has:
i) $f^t(x) \in C(x)$,*for all* $t \geq 1$;
ii) $\overline{x} \in C(x)$;
iii) *the following statements are equivalent:*
a) $x = \overline{x}$,
b) $C(x) = \{x\}$,
c) $x \in C(x)$.

Proof. *i)* $d(x, f^t(x)) \leq \sum_{k=0}^{t-1} d(f^k(x), f^{k+1}(x)) \leq \sum_{k=0}^{t-1} \alpha^k d(x, f(x)) = \frac{1-\alpha^t}{1-\alpha} d(x, f(x)) \leq \frac{1}{1-\alpha} d(x, f(x))$. Hence $d(x, \overline{x}) \leq \frac{1}{1-\alpha} d(x, f(x))$. For the same reason, since $\overline{x}$ is the fixed point of f^t, which is a contraction of constant α^t, one has

$$d(x, \overline{x}) \leq \frac{1}{1-\alpha^t} d(x, f^t(x)) \leq \frac{1}{1-\alpha} d(x, f^t(x)),$$

whence $d(x, f^t(x)) \geq (1-\alpha)\, d(x, \overline{x})$. On the other hand,

$$d(x, f(x)) - d(x, \overline{x}) \leq d(\overline{x}, f(x)) = d(f(\overline{x}), f(x)) \leq \alpha d(\overline{x}, x),$$

so that $d(x, f(x)) \leq (1+\alpha)\, d(x, \overline{x})$, i.e. $d(x, \overline{x}) \geq \frac{1}{1+\alpha} d(x, f(x))$. Thus we obtain $d(x, f^t(x)) \geq \frac{1-\alpha}{1+\alpha} d(x, f(x))$.

ii) It is an immediate consequence of *i)*.

iii) Implications *a)* $\Longrightarrow$ *b)* $\Longrightarrow$ *c)* $\Longrightarrow$ *a)* are obvious. ■

Corollary 12 *Let (X,d), f and a be as in Theorem 11. If the set A of acceptable solutions is the singleton of the fixed point of f, the mapping $N : X \to 2^X$ defined by $N(x) = X \setminus C(x)$ is a learning mapping of rejection for the decision problem $((X,d), A)$; moreover, the function r_N satisfies condition (1). Hence, if $(LLP)_N$ does not stop then every limit point of the sequence $\{x_t\}$ generated by $(LLP)_N$ is the fixed point of f; if, in addition, X is compact then $d(x_t, f(x_t)) \to 0$.*

Proof. The proof is similar to that of Thm. 9, using Thm. 3 and the inequality $r_N(x) \geq \frac{1-\alpha}{1+\alpha} d(x, f(x))$. ■

In the case when, as the preceding corollary suggests, we set $N(x) = X \setminus C(x)$, Theorem 11.*i)* shows that a possible choice in Step 2 of $(LLP)_N$, for $t > 0$, is $x_t = f(x_{t-1})$. In fact, under this choice the well known Banach contraction principle tells us that the sequence $\{x_t\}$ itself converges to the fixed point of f.

6 A Generalized "Tatonnement Process" for Best Reply Functions of Normal Form Games

In this paragraph we want to show that our learning process based on the principle of "Solutions Cannot Be There" is a generalized Cournot tatonnement (fictitious play) process for best replies functions of normal form games.

Consider a normal form game $G_i : (x_1, x_2, ..., x_n) \longmapsto G_i(x_1, x_2, ..., x_n) \in \mathbf{R}$, $i \in I = \{1, 2, ..., n\}$, where each player i chooses action $x_i \in X_i$ and gets the reward $G_i(x_1, x_2, ..., x_n)$. The set X_i is a complete metric space. Let $x = (x_1, x_2, ..., x_n) \in X = \Pi_{i=1}^n X_i$ denote the profile of the actions of all players. The best reply correspondence of this game is (assuming that it is well-defined)

$$\begin{array}{ll} X \to X, & \text{where} \\ x \longmapsto \Omega(x) = (\Omega_1(x), ..., \Omega_n(x)) & \end{array}$$

$$\Omega_i(x) = \arg\max_{x_i \in X_i} G_i(x_1, x_2, ..., x_{i-1}, x_i, x_{i+1}, ..., x_n).$$

If agents have common conjectures, or beliefs, over the actions of all other agents, which make vector $\widehat{x}$, then each of them realizes an action $x_i \in \Omega_i(\widehat{x})$ and makes the errors of anticipation, or of conjecture, $e_j(\widehat{x}) = d(\Omega_j(\widehat{x}), \widehat{x}_j)$, $j \in I \setminus \{i\}$.

A tatonement process or fictitious play (see Cournot [4], Moulin [17], Monderer and Shapley [16] and Foster and Young ([9]) works as follows: At period $t = 0$, the conjectures of the agents about their actions are $\widehat{x}(0)$. Their best replies generate their effective actions $x_i(0) = \Omega_i(\widehat{x}(0))$, $i \in I$. Agents realize that they have made errors of conjecture $e_i(\widehat{x}(0)) = d(\Omega_i(\widehat{x}(0)), \widehat{x}_i(0))$, $i \in I$.

At period $t = 1$, all agents make naive conjectures $\widehat{x}_i(1) = x_i(0) = \Omega_i(\widehat{x}(0))$, $i \in I$. Then the process repeats. It generates the new realized actions $x_i(1) = \Omega_i(\widehat{x}(1)) = \Omega_i^{(2)}(\widehat{x}(0))$, $i \in I$. The errors of conjecture are

$$e_i(\widehat{x}_i(1)) = d(\Omega_i(\widehat{x}(1)), \widehat{x}_i(1)) = d(\Omega_i^2(\widehat{x}(0)), \Omega_i(\widehat{x}(0))).$$

More generally,

$$e_i(\widehat{x}_i(t)) = d(\Omega_i^{t+1}(\widehat{x}(0)), \Omega_i^t(\widehat{x}(0))).$$

This process of conjectures due to Cournot [4] is curious, because it supposes that all agents do not take care of previous errors of conjecture to make their new conjectures! This is realistic enough only if the process of errors itself is strictly decreasing: $e_i(\widehat{x}_i(t+1)) < e_i(\widehat{x}_i(t))$, $t = 0, 1, 2....$; or $d(\Omega_i^{t+2}(\widehat{x}(0)), \Omega_i^{t+1}(\widehat{x}(0))) < d(\Omega_i^{t+1}(\widehat{x}(0)), \Omega_i^t(\widehat{x}(0)))$. This hypothesis is a bounded rationality assumption, based on the fact that agents realize perfectly that they make errors and deal with this reality to make less and less important errors.

Example 13 *Suppose that the best reply mappings Ω_i, $i \in I$, are contractions, i.e. $d(\Omega_i(x), \Omega_i(y)) < \alpha d(x, y)$ for all $x, y \in X$, $0 < \alpha < 1$. Then the process of* *errors of conjectures is such that* $d\left(\Omega_i(\widehat{x}), \Omega_i^2(\widehat{x})\right) < \alpha d\left(\widehat{x}, \Omega_i(\widehat{x})\right)$ *for all initial conjectures* $\widehat{x} \in X$. *The repetition of the process gives* $d\left(\Omega_i^t(\widehat{x}), \Omega_i^{t+1}(\widehat{x})\right) < \alpha^t d\left(\widehat{x}, \hat{\Omega_i}(\widehat{x})\right)$. *Then the inequalities* $e_i(\Omega_i^t(\widehat{x})) < \alpha^t e_i(\widehat{x})$, $i \in I$, *show that the process of errors of contracting best reply functions go to zero,for all initial conjectures* $\widehat{x}$.

The question is to know if the best reply mapping of a normal form game can define a “solutions cannot be there” dynamic process of learning that generalizes the tatonnement process (fictitious play dynamics) and converges towards the equilibrium. In this context, let $A = \{x \in X \ / \ \Omega_i(x) = x_i \ \forall i \in I\}$ denote the set of fixed points of the best reply mapping. We can say “yes” in the case of Lipschitzian best replies, because in this case we can exhibit a rejected set at each $x \in X$, as shown above. The “solutions cannot be there” process works as follows:

-Step 0. Set $t = 0$ and $F_t = X$.
-Step 1. If $F_t = \emptyset$, then $A = \emptyset$. Stop.
-Step 2. If $F_t \neq \emptyset$, choose $x_t = (x_{t1}, x_{t2}, ..., x_{tn}) \in F_t$.
-Step 3. If x_t is a fixed point of Ω, then $x_t \in A$. Stop.
-Step 4. If x_t is not a fixed point of Ω, i.e. $\Omega_i(x_t) \neq x_{ti}$ for some $i \in I$, construct a rejected ball $B\left(x_{ti}, r_i^{\Omega}(x_{ti})\right)$, with $r_i^{\Omega}(x_{ti}) > 0$.
-Step 5. Define $F_{t+1} = F_t \backslash R_t$, where $R_t = \Pi_{j \in I} R_{tj}$, $R_{tj} = X_j$ if $j \neq i$, $R_{ti} = B\left(x_{ti}, r_i^{\Omega}(x_{ti})\right)$. Increase t by 1 and go to Step 1.

The major improvement of our process (if we compare it with the tatonnement, or fictitious play, process) is that we are not obliged to choose the naive conjectures $\Omega_i(x_t)$ as the new conjectures. Lipschitzian best replies do not belong to the rejected balls, i.e. $\Omega(x) \notin B\left(x, r(x)\right)$. This shows that the best reply $\Omega(x)$ at point x is an acceptable conjecture. But, if $\Omega(x)$ is injective, this new conjecture $\Omega(x)$ cannot be a fixed point unless x itself is a fixed point. This shows the weakness of the naive (“a priori wrong”) conjectures of the Cournot tatonnement (or fictitious play) process. Our process is more general and is not contradictory.

In the usual tatonnement process, agents learn nothing except the fact that, if $\Omega(x_0) \neq x_0$, x_0 is not a fixed point of the best reply function and, if Ω is injective, all successors $\Omega^t(x_0)$ of x_0 cannot be fixed points. So they learn that their conjectures will be wrong at each stage! They do not learn some local information about the localization of the fixed points (balls where they cannot be), as in our process.

Adaptive models include fictitious play and other models. Monderer and Shapley [16], Gul [11] and Milgrom and Roberts [15] study partial or complete fictitious play, where only one player or all players, respectively, give their best replies at each period against the previous actions of all the other players. This poses different information problems.

7 Related Literature

What are the properties of the "Solutions Cannot Be There" process of learning, i.e. the process of successive elimination, period after period, of solutions not too far from non satisfactory present solutions? We are interested in contexts in which to repeat near an action that proved to be an error is an error, too. This is the case for normal form games on metric spaces, with Lipschitzian (in particular, contracting) best replies.This is also the case, but in a very different way, for increasing best replies and for best replies f such that $f(x) \geq x$.We will not examine this point here. Let us just say that if one knows that best replies satisfy the inequality $f(x) \geq x$ for all $x \in X$, where X is a CPO (an ordered set with a bottom element such that every nonempty chain in X has a supremun) and f is a mapping from X into itself then a fixed point of f cannot be located near x (more precisely, in the interval between x and $f(x)$). In this case, Davey and Priestly [6] show that f has a fixed point; this is ageneralization of Tarski fixed point theorem, which says that if X is a complete lattice and f is an order-preserving map then f has a fixed point. But in these cases we know much more: we can identify a fixed point, namely, the supremum of the subset of points x such that $f(x) \geq x$, in the case of Tarski theorem, or the limit point of a chain $f^t(x)$, $t = 1, 2, 3, ...$, in the more general case.

Elimination of dominated strategies can also be interpreted as a very strong form of our "Solutions Cannot Be There" principle.

Following Sobel [20], learning models can be classified in three ways, depending on how agents collect information, the strategic aspect of their environment and their degree of rationality. Our paper assumes that agents collect informations by doing actions. This is yet the essence of the two armed bandit model [18], where information is costly because the only way to know that the good choice has been made is choosing some action. "An agent will not necessarily learn the optimal decision when the cost of acquiring additional information exceeds the benefit"[20]. So the agent can be locked into a particular bad strategy. The Bandit Problem assumes, as we do, that the environment is stationary. But we do not get this unpleasant locking aspect, which is wrong in a changing environment. Two concepts appear to be very important: how agents make errors and how rapidly they correct them.(the rapidity of learning).

Errors come into the picture because agents lack information, do not know the payoffs of other players and have limited understanding of the problems they try to solve. Agents make bad decisions because they are inexperienced, ignorant or face difficult problems Models must necessarily specify what agents initially know and how they build this knowledge. In the case of herding, Smith and Sorensen[2] model situations in which agents can make bad decisions and, after a finite number of periods, agents repeatedly make the same choice. The main process consists in essays, errors and corrections of errors. Agents can also make goods decisions when they have experienced making similar decisions.

[2] "Informational Herding and Optimal Experimentation", Department of Economics Working Paper 97/22, MIT, 1997.

Secondly, the rapidity of learning, i.e. how fast agents learn, is very important for practical reasons. The Engineering literature emphasizes much more the shape of the learning curve [5]. The great problem is to know when does learning lead to good results in a realistic amount of time. Agents learn by using past experiences to obtain a better understanding of their current and future problems and to improve their decisions, identifying circumstances under which learning leads to optimal decisions. Learning can be complete in finite time only if the environment is stationary. We agree with Aghion, Bolton, Harris and Jullien [1], which show, on the contrary, that complete learning requires strong assumptions. They identify situations where agents learn the correct decision asymptotically.

The strategic aspect of learning can emerge when there are several agents. We deal with this aspect when the acceptable set is the set of equilibria of a normal form game. Agents must observe the actions done by others and must be sure that, if an equilibrium is not attained, the equilibrium cannot be very close to the profile of actions just taken.

The degree of rationality we attribute to agents is very low. We agree with the observation of Sobel [20] that "the more complicated the strategic environment, the simpler the information technology and the simpler the behaviour followed by the agents". In our model agents do not necessarily optimize. They can follow behavioural rules. We can suppose bounded rationality and satisfying:agents who do not know where the set of acceptable solutions is located. Our model is a kind of search model of acceptable points. Satisfying is a process by which agents approach some acceptable solutions (like equilibrium); it works by elimination, like the brain does (connectionist models). Information comes, each period, from the actions taken by all players just before; the only way an agent can find out how well a choice works is by making the choice.

More generally, we can relate our learning process to the existing literature on Learning following Sobel [20], who makes clear that three approaches emerge as a good starting point for learning models:

i) Individual learning models, like the celebrated two armed bandit problem [18], and "models of experiments" of Aghion, Bolton, Harris and Jullien [1].

ii) Social learning models, which include observational learning models with the possibility of inefficient herding [?]. Models of word-of-mouth transmission of information between boundedly rational agents examine how agents learn. The most prominent model of this kind is that of A.V. Banerjee and D. Fudenberg[3], followed by Ellison and Fudenberg [8], which makes the asymptotic behaviour of agents precise. Smallwood and Conslik [19] demonstrate that the rationality of agents change the qualitative results of the model.

iii) Evolutionary and learning models in games include learning models that examine fictitious play and tatonnement (adaptive) dynamic processes, where learning may be slow [10], [16], and evolutionary game models, close to models of social learning, which describe non strategic agents that can mutate with some probability and change actions infrequently. Recently, Kandori, Mailath and

[3] "Word-of-Mouth Learning", mimeo, MIT and Harvard University, 1995.

Rob [12] and Foster and Young [9] have examined strategic adaptive processes where agents adjust their strategies by playing best responses to the previous population distribution with a positive probability of a random change in their strategies.

8 Conclusions and Future Work

In this paper we have proposed a new principle of learning, the "Solutions Cannot Be There" principle. We have presented a model in which agents learn from errors and we have proved that, under mild assumptions, their actions converge to acceptable solutions of their decision problems. In the specific case when these acceptable solutions are the fixed points of a continuous mapping from a metric space into itself, we have proposed a way to link positively the amount of learning to the size of errors. As possible topics for further research, one could consider extending our model to nonstationary environments and to introduce decision costs [3] for measuring the difficulty of making decisions, linking these costs to the sizes of the rejected sets.

References

[1] P. Aghion, P. Bolton, C. Harris and B. Jullien, Optimal Learning by Experimentation, *Rev. Econ. Stud.* 58 (1991), 621-654.

[2] J.-P. Aubin, "Applied Abstract Analysis," John Wiley & Sons, New York, 1977.

[3] J. Conlisk, Why Bounded Rationality?, *J. Econ. Lit.* 34 (1996), 669-700.

[4] A. Cournot, "Recherches sur les Principes Mathématiques de la Théorie de la Richesse," Hachette, Paris, 1838.

[5] E. Dar-El, "Human Learning: From Learning Curves to Learning Organizations", Kluwer Academic Publishers, Boston, 2000.

[6] B. A.Davey and H. A. Priestly, "Introduction to Lattices and Order," Cambridge University Press, Cambridge, 1990.

[7] G. De Marco, For Every ε There Continuously Exists a δ, *Am. Math. Mon.* 108 (2001), 443-444.

[8] G. Ellison and D. Fudenberg, Word-of-Mouth Communication and Social Learning, *Quart. J. Econ.* 110 (1995), 93-125.

[9] D. Foster and P. Young, On the Convergence of Fictitious Play in Coordination Games, *Games Econ. Behaviour* 25 (1998), 79-96.

[10] D. Gale, What We Have Learned from Social Learning, *Europ. Econ. Rev.* 40 (1996), 617-628.

[11] F. Gul, Rationality and Coherent Theories of Strategic Behavior, *J. Econ. Theory* 70 (1996), 1-31.

[12] M. Kandori, G. Mailath and R. Rob, Learning to Play Equilibria in Games with Stochastic Pertubations, *Econometrica* 61.(1993), 29-56.

[13] N. V. Long and A. Soubeyran, Existence and Uniqueness of Cournot Equilibrium: A Contraction Mapping Approach, *Economics Letters* 67 (2000), 345-348.

[14] W. R. Mann, Mean Value Methods in Iteration, *Proc. Am. Math. Soc.* 4 (1953), 506-510.

[15] P. Milgrom and J. Roberts, Adaptive and Sophisticated Learning in Normal Form Games, *Games Econ. Behavior* 3 (1991), 82-100.

[16] D. Monderer and L. S. Shapley, Fictitious Play Property for Games with Identical Interests, *J. Econ. Theory*, 68 (1996), 258-265.

[17] H. Moulin, Dominance Solvability and Cournot Stability, *Math. Social Sci.* 7 (1984), 83-102.

[18] M. Rothschild, A Two Armed Bandit Theory of Market Pricing, *J. Econ. Theory* 9 (1974), 185-202.

[19] D. Smallwood and J. Conlisk, Product Quality in Markets where Consumers are Imperfectly Informed, *Quart. J. Econ.* 93 (1993), 1-23.

[20] J. Sobel, Economists' Models of Learning, *J. Econom. Theory* 94 (2000), 241-261.