Generalized Inexact Proximal Algorithms: Habit's/ Routine's Formation with Resistance to Change, following Worthwhile Changes

G. C. Bento* A. Soubeyran †

April 1, 2014

Abstract

This paper shows how, in a quasi metric space, an inexact proximal algorithm with a generalized perturbation term appears to be a nice tool for Behavioral Sciences (Psychology, Economics, Management, Game theory,...). More precisely, the new perturbation term represents an index of resistance to change, defined as a "curved enough" function of the quasi distance between two successive iterates. Using this behavioral point of view, the present paper shows how such a generalized inexact proximal algorithm can modelize the formation of habits and routines in a striking way. This idea comes from a recent "variational rationality approach" of human behavior which links a lot of different theories of stability (habits, routines, equilibrium, traps,...) and changes (creations, innovations, learning and destructions,...) in Behavioral Sciences and a lot of concepts and algorithms in Variational Analysis. In

*IME-Universidade Federal de Goiás, Goiânia-GO 74001-970, Caixa Postal: 131, BR,

Fone: +55 (62)3521-1208 (glaydstonc@gmail.com)

†Aix-Marseille School of Economics, Aix-Marseille University, CNRS & EHESS, FR (antoine.soubeyran@gmail.com)

this variational context, the perturbation term represents a specific instance of the very general concept of resistance to change, which is the disutility of some inconvenients to change. Central to the analysis are the original variational concepts of "worthwhile changes" and "marginal worthwhile stays". At the behavioral level, this paper advocates that proximal algorithms are well suited to modelize the emergence of habituation/routinized human behaviors. We show when, and at which speed, a "worthwhile to change" process converges to a behavioral trap.

AMS Classification. 49J52 · 49M37 · 65K10 · 90C30 · 91E10.

keywords. Nonconvex optimizationKurdyka-Lojasiewicz inequalityinexact proximal algorithmshabitsroutinesworthwhile changes

1 Introduction

The main message of this paper is that, using the behavioral context of a recent "Variational rationality" approach of worthwhile stay and change dynamics proposed by Soubeyran [1, 2], a generalized proximal algorithm can modelize fairly well an habituation process as described in Psychology for an agent, or a routinization process, in Management Sciences, for an organization. This opens the door to a new vision of proximal algorithms. They are not only very nice mathematical tools in optimization theory, with striking computational aspects. They can also be nice tools to modelize the dynamics of human behaviors.

Theories of stability and change consider successions of stays and changes. Stays refer to habits, routines, equilibrium, traps, rules, conventions,. . . . Changes represent creations, destructions, learning processes, innovations, attitudes as well as beliefs formation and revision, self regulation problems, including goal setting, goal striving and goal revision, the formation and break of habits and routines,. . . . In the interdisciplinary context which characterizes all these theories in Behavioral Sciences, the "Variational rationality approach" (see [1, 2]), shows how to modelize the course of human activities as a succession of worthwhile temporary stays and changes which balance, each step, motivation to change (the utility of advantages to change) and resistance to change (the disutility of inconvenients to change). This very simple idea has allowed to see proximal algorithms as an important tool to modelize the human course of actions, where the perturbation term of a proximal algorithm can be seen as a crude formulation of the complex concept of resistance to change, while the utility generated by a change in the objective function can represent a crude formulation of the motivation to change concept. The Variational rationality approach considers three original concepts: i) "worthwhile changes", when, each step, motivation to change is higher enough with respect to resistance to change, ii) "non marginal worthwhile changes" and, iii) "variational traps", "easy enough to reach", that the agent can reach using a succession of worthwhile changes, and "difficult enough to leave", such that, being there, it is not worthwhile to move from there.

These three concepts represent the pilar of the Variational rationality approach; see [1, 2], which has provided an extra motivation to develop further the study of proximal algorithms in a nonconvex and possibly nonsmooth setting. Among other recent applications of this simple idea, see Attouch and Soubeyran [3] for local search proximal algorithms, Flores-Bazan et al. [4] for worthwhile to change games, Attouch et al. [5] for alternating inertial games with costs to move, and Cruz Neto et al. [6] for the "how to play Nash" problem,.... In all these papers the perturbation term of the usual proximal point algorithm is a linear or a quadratic function of the distance or quasi distance between two successive iterates. They modelize the case of "strong enough resistance" to change. Our paper examines the opposite case of "weak enough" resistance to change where the pertubation term modelizes the difficulty (relative resistance) to be able to change as a "curved enough" function of the quasi distance between two successive iterates. A quasi distance modelizes costs to be able to change as an index of dissimilarity between actions where the cost to be able to change from an action to an other one is not the same as the cost to be able to change in the other way. In a first paper, Bento and Soubeyran [7] show when, in a quasi metric space, a generalized inexact proximal algorithm, equipped with such a generalized perturbation term, defined, each step, by a sufficient descent condition and a stopping rule, converges to a critical point. Then, it is shown that the speed of convergence and convergence in finite time depend of the curvature of the perturbation term and of the Kurdyka-Lojasiewicz property associated to the objective function. A striking and new application is given. It concerns the impact of the famous "loss aversion effect" (Nobel Prize Kahneman and Tversky [8], Tversky and Kahneman [9]) on the speed of convergence of the generalized inexact proximal algorithm.

In the present paper, inspired by the VR "variational rationality" approach, we consider a new inexact proximal algorithm whose sufficient descent condition is, each step, a little more demanding, using, each step, the same stopping rule. Applying the convergent result of the first paper [7] this simple modification is a way to force convergence even more. It gives an intuitive sufficient condition for the critical point to be a variational trap (weak or strong). In this case changes are required to be " worthwhile enough", the stopping rule is the same, and the end of the convergent worthwhile stay and change process is both a critical point and a variational trap. Doing so, this paper extends the convergence result to a critical point of Attouch and Bolte [10], Attouch et al. [11] and Moreno et al. [12], using a fairly general "convex enough" perturbation term. It is important to note that, as an application, it is possible to consider the formation of habits and routines as an inexact proximal algorithm in the context of weak resistance to change. However, because of its strongly interdisciplinary aspect (Mathematics, Psychology, Economics, Management), to be carefully justified, this application needs several steps. Due to space constraints, these considerations are given in Bento and Antoine [13]. Then, at the behavioral level, the main message of this paper is to advocate that our generalized proximal algorithm is well suited to modelize the formation of habitual/ routinized human behaviors. The list of the main (VR) concepts is presented on an example in Section 2. To get more perspective, in the Annex, the VR approach of stability and change dynamics is compared in great detail with a complementary theory relative to the dynamics of human behavior, the HD habitual domain theory (see Yu [14]) and its recent continuation, the DMCS approach (decision making with changeable spaces; see Larbani and Yu [15]). A last section applies the convergence and speed of convergence result of our generalized proximal algorithm to the modelization of the formation and break of habits and routines. It compares very succinctly what HD and VR can say on this important topic for human behavior.

At a higher dimensional level where inexact proximal algorithms represent a specific formulation of VR, the complementary VR and HD approaches. consider, both stability and change dynamics, but modelize them in a different way: deterministic worthwhile temporary stays and changes dynamics for the VR approach and Markov chains for HD theory. Both focus the attention on optimization and satisficing processes. VR is variable and possibly intransitive preference (utility) based, while HD is charge based. VR main topic is the self regulation problem (goal setting, goal striving, goal revision and goal disengagement) at the individual level or for interactive agents. The HD main topic is to know "how agents expand and enrich their habitual domain". Both approaches examine decision making problems with changeable structures (spaces, parameters, goals, preferences and charges). Each approach considers different aspects of what can be changed and how it can change.

Our paper is organized as follows. Section 2 gives an example which helps to list the main variational tools necessary to define the central concepts of "worthwhile change" and "variational trap" for behavioral applications. Section 3 shows how inexact proximal algorithms can represent adaptive satisficing processes. Section 4 examines a generalized inexact proximal algorithm which converges to a critical point which is also a variational trap (weak or strong), when the objective function satisfies a Kurdyka-Lojasiewicz inequality. A last section summarizes very briefly the VR variational rationality approach and the HD habitual domain theory. The conclusion follows. In [13] the authors compare in greater details the VR and HD approaches relative to habituation/routinization processes.

2 Variational Rationality : How Successions of Worthwhile Stays and Changes End in Variational Traps

2.1 Worthwhile Stay and Change Dynamics

A recent variational rationality approach, see [1, 2], gives a common background to a lot of theories of stability/stay and change in Behavioral Sciences (Psychology, Economics, Management Sciences, Decision theory, Philosophy, Game theory, Political Sciences, Artifical Intelligence...), using as a central building bloc the three concepts of "worthwhile change", "marginal worthwhile change" and "variational trap". All these behavioral dynamics can be seen as a succession of worthwhile temporary stays and changes $x^{k+1} \in W_{e_k,\xi_{k+1}}(x^k)$, $k \in \mathbb{N}$, ending in variational traps $x^* \in X$, where X is the universal space of actions (doing), having or being, depending of the applications. X includes all past elements and all the new elements that can be discovered as time evolves.

The main idea is quite evident. If a behavioral theory wants to explain "why, where, how and when" agents perform actions and change, this theory must define, each period, along a path of changes $\{x^0, x^1, ..., x^k, x^{k+1}, ...\}$ why the agent have, first, an incentive to do some steps away from his current position and, then, an incentive to stop changing one step more within this period. In the current period $k+1$, a change is such that $x^{k+1} \neq x^k$, while a stay is $x^{k+1} = x^k$. Let $e_k \in E$ be the experience of the agent at the end of the last period k. A change $x^k \curvearrowright x^{k+1} \in W_{e_k,\xi_{k+1}}(x^k)$ is worthwhile, when his ex ante motivation to change $M_{e_k}(x^k, x^{k+1})$ is sufficiently higher (more than $\xi_{k+1} > 0$) than his ex ante resistance to change, $R_{e_k}(x^k, x^{k+1})$. Then, $x^{k+1} \in W_{e_k,\xi_{k+1}}(x^k) \Longleftrightarrow M_{e_k}(x^k, x^{k+1}) \geq \xi_{k+1} R_{e_k}(x^k, x^{k+1})$. Motivation and resistance to change are two complex variational concepts which admit a lot of variants (see [1, 2]). Motivation to change $M_{e_k}(x^k, x^{k+1}) = U_{e_k}\left[A_{a_k}(x^k, x^{k+1})\right]$ is the utility $U_{e_k}[\cdot]$ of advantages to change, $A_{e_k}(x^k, x^{k+1})$, while resistance to change $R_{e_k}(x^k, x^{k+1}) = D_{e_k}\left[I_{e_k}(x^k, x^{k+1})\right]$ is the disutility $D_{e_k}[\cdot]$ of inconvenients to change $I_{e_k}(x^k, x^{k+1})$.

Worthwhile changes are generalized satisficing changes: Within a period, a worthwhile change $x^k \curvearrowright x^{k+1} \in W_{e_k,\xi_{k+1}}(x^k)$ is desirable and feasible enough, i.e, acceptable, improving with no too high costs to be able to improve. Then, a worthwhile change is a generalized satisficing change where, each period, the agent chooses the ratio $\xi_{k+1} > 0$, which represents how worthwhile a change must be to accept to move rather than to stay. The famous Simon [16] satisficing principle is a specific case (see [1, 2]). Second, within the same period, the agent must also have to know when he must stop changing. This is the case when one step more is not worthwhile. More formally, this change is not "marginally worthwhile", when the ex ante marginal motivation to change is sufficiently lower than the ex ante marginal resistance to change. In this case the agent does not regret ex ante to do not go one step further. The motivation to change again next period comes from residual unsatisfied needs or variable preferences.

A variational trap x^* is such that, starting from an initial point $x^0 \in X$, it exists a path of worthwhile changes $x^{k+1} \in W_{e_k,\xi_{k+1}}(x^k)$ which ends in x^*, i.e., such that, being there, it is not worthwhile to move again, i.e., $W_{e_*,\xi_*}(x^*) = \{x^*\}$.

2.2 Variational Concepts. An Example

To save space and to fix ideas, let us define on a simple example all these variational rationality concepts. This being done, we can easily show how an inexact proximal algorithm represents a nice benchmark process of worthwhile temporary stays and changes in term of, each period, a sufficient descent condition and a stopping rule. For much more comments, and a more complete formulation of each of these variational concepts, with references to a lot of different disciplines in Behavioral Sciences which help to justify their unifying power; see [1, 2].

A simple model of knowledge management: This example modelizes a very simple case of knowledge management within an organization, to determine a satisficing or, as a extreme case, the optimal size and shape of an innovative firm driven by a leader. In Management Sciences, the literature on this topic is enormous and represents one of its main area of research. Consider an entrepreneur (leader) who, each period, can hire and fire different kinds and numbers of skilled and specialized workers $\{1, 2, ..., j, \ldots, l\} = J$ (say knowledge workers; see Long et al. [17]) to produce a chosen quantity of a final good of a chosen quality. The endogenous quality $q(x)$ of this final good changes with the chosen profile of skilled workers $x = (x^1, x^2, .., x^j, .., x^l) \geq 0$, where $x^j \geq 0$ is a number of workers of type j. To save space and for simplification, each period, each employed skilled worker utilizes one unit of a specific non durable mean to produce, using his specific knows how to produce, one unit of a specific component of type j. Then, the entrepreneur combines these different components to produce $\mathfrak{q}(x)$ units of a final good of endogenous quality $s(x)$. This production function is original because it mixes both variable quantity and quality. The revenue of the entrepreneur is $\varphi\left[\mathfrak{q}(x), s(x)\right]$. His operational costs $\rho(x)$ are the sum of his costs to buy the non durable means used by each worker, and the wages paid to each employed worker. Then, in a given period, the profit of the entrepreneur who employs the profile $x \in X = R^l$ of skilled workers is $g(x) = \varphi\left[\mathfrak{q}(x), s(x)\right] - \rho(x) \in R$. For a famous example of an endogenous production function of quality, see Kremer [18].

Advantages to change: Let $x = x^k$ and $y = x^{k+1}$ be the last period, and current period profiles of skilled workers chosen by the entrepreneur. Then, if this is the case, his advantages to change his profile of skilled workers from one period to the next is $A(x, y) = g(y) - g(x) \geq 0$;

Inconvenients to change: They represent the difference $I(x, y) = C(x, y) - C(x, x) \geq 0$ between costs $C(x, y)$ to be able to change from profile x to profile y and costs $C(x, x)$ to be able to stay with the same profile x used in the last period;

Costs to be able to change (to stay): To be able to hire one skilled worker of type j, ready to work, costs $h_+^j > 0$. These costs include search and training costs. To fire one worker of type j, costs $h_-^j > 0$. These costs represent separation and compensation costs. To keep a worker, ready to work, one period more, costs $h_=^j \geq 0$. These conservation costs include knowledge regeneration and motivation costs. Then, in the current period, i) costs to conserve the same profile of workers as in the last period are $C(x,x) = \Sigma_{j=1}^n h_=^j x^j$ while , ii) costs to utilize the profile of skilled workers y are:

$$C(x,y) = \Sigma_{j\in J_+(x,y)} \left[h_=^j x^j + h_+^j (y^j - x^j)\right] + \Sigma_{j\in J_-(x,y)} \left[h_=^j y^j + h_+^j (x^j - y^j)\right],$$

where $J_+(x,y) = \left\{j \in J, y^j \geq x^j\right\}$ and $J_-(x,y) = \left\{j \in J, y^j < x^j\right\}$. For simplification, suppose that conservation costs are zero, i.e., $h_=^j = 0$. Then,

$$C(x,x) = 0 \quad \text{and} \quad I(x,y) = \Sigma_{j\in J_+(x,y)} h_+^j (y^j - x^j) + \Sigma_{j\in J_-(x,y)} h_+^j (x^j - y^j).$$

So, $I(x,y)$ is a quasi distance $q(x,y) := I(x,y) \geq 0$ such that

i) $q(x,y) = 0$ iff $y = x$;

ii) $q(x,z) \leq q(x,y) + q(y,z)$, $x,y,z \in X$.

The more general case where $h_=^j > 0$ works as well.

Motivation and resistance to change functions: They are, moving from the past profile of knowledge workers x to the current profile y are

$$M(x,y) = U\left[A(x,y)\right] = \left[g(y) - g(x)\right]^\mu \quad \text{and} \quad R(x,y) = D\left[I(x,y)\right] = q(x,y)^\nu, \quad \mu,\nu > 0,$$

where the utility and disutility functions are $U[A] = A^\mu$ and $D[I] = I^\nu$.

Relative resistance to change function: It is $\Gamma\left[q(x,y)\right] = U^{-1}\left[D\left[I(x,y)\right]\right] = q(x,y)^{\nu/\mu}$, where $\alpha = \nu/\mu > 0$.

Worthwhile changes: In this setting, a change from profile x to profile y is worthwhile if $M(x, y) \geq \xi R(x, y)$, i.e., $[g(y) - g(x)]^{\mu} \geq \xi q(x, y)^{\nu}$, where $\xi > 0$ is the current and chosen "worthwhile enough" satisficing ratio. Then, a worthwhile change is such that

$$y \in W_{\xi}(x) \Longleftrightarrow g(y) - g(x) \geq \lambda \Gamma \left[q(x, y) \right], \qquad \lambda = (\xi)^{1/\mu} > 0.$$

Succession of worthwhile temporary stays and changes: In this example they are

$$g(x^{k+1}) - g(x^k) \geq \lambda_{k+1} \Gamma \left[q(x^k, x^{k+1}) \right], \qquad k \in \mathbb{N}.$$

Variational traps: In the example, given the initial profile of skilled workers $x^0 \in X$, and a final worthwhile enough to change ratio $\lambda_* > 0$, $x^* \in X$ is a variational trap if it exists a path of worthwhile temporary stays and changes $\left\{ x^0, x^1, ..., x^k, x^{k+1}, ... \right\}$ such that,

i) $g(x^{k+1}) - g(x^k) \geq \lambda_{k+1} \Gamma \left[q(x^k, x^{k+1}) \right]$, $k \in \mathbb{N}$;

ii) $g(y) - g(x^*) < \lambda_* \Gamma \left[q(x^k, y) \right]$, $y \neq x^*, y \in X$.

An habituation/routinization process: It is such that, step by step, gradually, the agent carries out a more and more similar action. This is equivalent to say than the quasi distance $C(x^k, x^{k+1})$ converges to zero as k goes to infinite.

When a worthwhile to change process converges to a variational trap, this variational formulation offers a model of trap as the end point of a path of worthwhile changes.

3 Inexact Proximal Algorithms as Worthwhile Stays and Changes Processes

3.1 Inexact Proximal Formulation of Worthwhile Changes

Proximal intransitive preferences. Let us define, in the current period $k+1$, the "to be increased" entrepreneur proximal payoff to change from $x = x^k$ to $y = x^{k+1}$ as $Q_\lambda(x, y) = g(y) - \lambda\Gamma\left[q(x, y)\right]$ with $\lambda > 0$. Then, the proximal payoff to stay at $x = x^k = y = x^{k+1}$ is $Q_\lambda(x, x) = g(x) - \lambda\Gamma\left[q(x, x)\right] = g(x)$. It follows that it is worthwhile to change from profile x to profile y iff $Q_\lambda(x, y) \geq Q_\lambda(x, x)$, i.e., $y \in W_\lambda(x)$. This defines a variable and possibly non transitive preference $z \geq_{x,\lambda} y \Longleftrightarrow Q_\lambda(x, z) \geq Q_\lambda(x, y)$. To fit with the formulation of inexact proximal algorithms, where mathematicians consider "to be decreased" cost functions, let us consider the residual profit that the entrepreneur expects to exhaust in the future, $f(x) = \overline{g} - g(x) \geq 0$, where $\overline{g} = \sup\left\{g(y), y \in X\right\} < +\infty$ is the highest finite profit that the entrepreneur can hope to get. Then, the "to be decreased" proximal payoff of the entrepreneur is

$$P_\lambda(x, y) = f(y) + \lambda\Gamma\left[q(x, y)\right]. \tag{1}$$

In this case, to move from profile x to profile y is a worthwhile change $y \in W_\lambda(x)$ iff $P_\lambda(x, y) \leq P_\lambda(x, x)$.

Sufficient descent methods. The entrepreneur performs, each period $k+1$, a sufficient descent, if he can choose a new profile x^{k+1} such that $f(x^k) - f(x^{k+1}) \geq \lambda_{k+1}\Gamma[q(x^k, x^{k+1})]$. This means that the entrepreneur follows a path of worthwhile changes $x^{k+1} \in W_{\lambda_{k+1}}(x^k)$, $k \in \mathbb{N}$. Since $q(x^k, x^k) = 0$, this comes from definition of $W_{\lambda_{k+1}}(x^k)$ combined with (1) for $x = x^k$, $y = x^{k+1}$ and $\lambda = \lambda_{k+1}$. In this case, each worthwhile change is not optimizing, contrary to each step of an exact proximal algorithm.

Exact proximal algorithms. The entrepreneur follows an exact proximal algorithm if, each current period $k+1$, he can choose a new profile x^{k+1} which minimizes his "to be decreased" proximal payoff

$P_{\lambda_{k+1}}(x^k, y) = f(y) + \lambda_{k+1}\Gamma[q(x^k, y)]$ on the whole space X,

$$x^{k+1} \in \text{argmin}_{y \in X} \left\{ f(y) + \lambda_{k+1}\Gamma[q(x^k, y)] \right\}, \qquad k \in \mathbb{N}, \tag{2}$$

which allows us to obtain $x^{k+1} \in W_{\lambda_{k+1}}(x^k)$, $k \in \mathbb{N}$. In Mathematics the formulation is

$$x^{k+1} \in \text{argmin}_{y \in X} \left\{ f(y) + \lambda_k\Gamma[q(x^k, y)] \right\}, \qquad k \in \mathbb{N}. \tag{3}$$

It takes λ_k instead of λ_{k+1}. In this case, the entrepreneur follows a path of optimal worthwhile changes, $x^{k+1} \in W_{\lambda_k}(x^k)$, $k \in \mathbb{N}$. In this paper, we will adopt the Mathematical formulation.

Epsilon inexact proximal algorithms. Several variants can be founded in this important literature of what is inexact. Let us consider the version given in Attouch and Soubeyran [3] following a long tradition, starting with Rockafellar [19]. In our context, the entrepreneur follows an inexact proximal algorithm if, each period $k+1$, he can choose a new profile x^{k+1} such that

$$f(x^{k+1}) + \lambda_k\Gamma[q(x^k, x^{k+1})] \le f(y) + \lambda_k\Gamma[q(x^k, y)] + \varepsilon_k, \quad y \in X,$$

given a sequence of nonnegative error terms $\{\varepsilon_k\}$, i.e, $P_{\lambda_k}(x^k, x^{k+1}) \le P_{\lambda_k}(x^k, y) + \varepsilon_k$, $y \in X$. The term λ_k can be replaced by λ_{k+1}.

Epsilon inexact proximal algorithms represent a succession of adaptive satisficing processes. Let $\overline{Q}_{\lambda_k}(x^k) = \sup\left\{Q_{\lambda_k}(x^k, y), y \in X\right\} < +\infty$ and $\underline{P}_{\lambda_k}(x^k) = \inf\left\{P_{\lambda_k}(x^k, y), y \in X\right\} > -\infty$ be, for each current period $k+1$, the optimal past values of the "to be increased" and 'to be decreased" proximal payoffs of this entrepreneur. Let $\overline{Q}_{\lambda_k}(x^k) - s_{k+1}$ and $\underline{P}_{\lambda_k}(x^k) + s_{k+1}$ be, in this current period $k+1$, the current satisficing levels of the "to be increased" and "to be decreased" proximal payoffs of the entrepreneur. In this current period, $s_{k+1} > 0$ represents, for the VR approach, a given satisfacing rate; see [1, 2]. For an inexact proximal algorithm, $s_{k+1} = \varepsilon_k > 0$ is a given error term.

Then, in the context of the VR theory, an inexact proximal algorithm have a new interpretation. It means that, for each period $k+1$, the new profile x^{k+1} must be satisficing. That is to say, "to be increased" and "to be decreased" proximal payoffs of the entrepreneur must be higher or lower than the current satisficing level, i.e., $Q_{\lambda_k}(x^k, x^{k+1}) \geq \overline{Q}_{\lambda_k}(x^k) - \varepsilon_k$ for a "to be increased" proximal payoff and $P_{\lambda_k}(x^k, x^{k+1}) \leq \underline{P}_{\lambda_k}(x^k) + \varepsilon_k$ for a "to be decreased" proximal payoff. Each period $k+1$, let us consider the variable satisficing set $S_{\lambda_k,\varepsilon_k}(x^k) = \left\{y \in X, P_{\lambda_k}(x^k, y) \leq \underline{P}_{\lambda_k}(x^k) + \varepsilon_k\right\}$. Then, an epsilon inexact proximal algorithm is defined by a succession of repeated decision making problems with changeable spaces and goals (satisficing levels): find $y \in S_{\lambda_k,\varepsilon_k}(x^k)$, $k \in \mathbb{N}$. They are decision making problems with changeable spaces. See Larbani and Yu [15] for different aspects of what can be change and how (their DMCS approach).

3.2 Marginally Worthwhile Changes

Consider the current period $k+1$. Let $x = x^k \curvearrowright y = x^{k+1}$ be a worthwhile change from x^k to $x^{k+1} \in W_{\lambda_k}(x^k)$ and let $x^{k+1} \curvearrowright z \in \mathfrak{M}(x^{k+1}) \subset X$ be a marginal change, where $\mathfrak{M}(x^{k+1})$ is a small neighborhood of x^{k+1} in the quasi-metric space X. Then, at each period $k+1$, the agent who has done the worthwhile change $y = x^{k+1} \in W_{\lambda_k}(x^k)$ will stop to prolonge this change if, doing one step more this period $k+1$, from x^{k+1} to $z \in \mathfrak{M}(x^{k+1})$, this marginal change is not worthwhile, i.e., $z \notin W_{\lambda_k}(x^{k+1})$. This is a generalized stopping rule, a "not worthwhile marginal change" condition, that will be used later in the context of proximal algorithms; see condition (12).

3.3 Classification of Inexact Proximal Algorithms: The Separation Between Weak and Strong Resistance to Change

Two cases. The consideration of relative resistance to change functions $\Gamma[\cdot]$ helps to classify proximal algorithms in two separate groups. The first case is of strong resistance to change, where $\Gamma[q] = q$ for all $q \geq 0$. This case have been examined in [1, 2, 3]. The second case is of weak resistance to change, where $\Gamma[q] = q^2$ and $q = q(x, y)$ is a distance and not a quasi distance. This is the traditional case. The literature on this topic is enormous; see, for example, Moreau [20] and Martinet [21], as well as in the study of variational inequalities associated to maximal monotone operators; see Rockafellar [19].

The variational approach which considers relative resistance to change as a core concept which balances motivation and resistance to change provides us an extra motivation to develop further the study of proximal algorithms in a nonconvex and possibly nonsmooth setting where the perturbation term of the usual proximal point algorithm becomes a "curved enough" function of the quasi distance between two successive iterates. Soubeyran [1, 2] and, later, Bento and Soubeyran [7], in a first paper which paves the way for the present one, have shown the strong link between a relative resistance to change index with the famous "loss aversion" index ([8, 9]). The generalized proximal algorithm examined, both, in [7] and in the present paper, is new and more adapted for applications in Behavioral Sciences. Moreover, it retrieves recent approaches of the proximal method for nonconvex functions; see [10, 12].

Hypothesis on the relative resistance to change. In the remainder of this paper we assume that Γ is a twice differentiable function such that:

$$\Gamma[0] = \Gamma'[0] = 0, \quad \text{and} \quad \Gamma'[q] > 0, \quad \Gamma''[q] > 0, \quad q > 0, \tag{4}$$

and there exist constants $r, \bar{q}, \bar{\rho}_\Gamma(r) > 0$, satisfying the following condition:

$$\Gamma'[q/r] \leq \bar{\rho}_\Gamma(r)\Gamma[q]/q, \qquad 0 < q \leq \bar{q}. \tag{5}$$

Let us consider a generalized rate of curvature of Γ given by:

$$\rho_\Gamma(q,r) := \frac{\Gamma'[q/r]}{(\Gamma[q]/q)}, \qquad 0 < q \leq \bar{q}. \tag{6}$$

In the particular case $r = 1$, (6) represents, in Economics, the elasticity of the disutility curve Γ; see, for instance, [1, 2]. From (6), condition (5) is equivalent to the condition:

$$\bar{\rho}_\Gamma(r) = \sup\{\rho_\Gamma(q,r) : 0 < q < \bar{q}\} < +\infty, \qquad r \in]0,1[\quad \text{fixed.}$$

Let us consider, for each $\alpha > 1$ fixed, the function $\Gamma[q] := q^\alpha$. It is easy to see that, in this case, $\bar{\rho}_D(r) \in [\alpha r^{1-\alpha}, +\infty)$. In particular, we can take

$$\bar{\rho}_\Gamma(q,r) = \alpha r^{1-\alpha} = \bar{\rho}_\Gamma(r) < +\infty. \tag{7}$$

More accurately, for each $\alpha > 1$, $\Gamma[q] = q^\alpha$ represents a disutility of costs to change. It is strictly increasing and satisfies (4) and (5).

4 An Inexact Proximal Point Algorithm: Convergence to a Weak or Strong Variational Trap

4.1 End Points as Critical Points or Variational Traps

In a first paper, Bento and Soubeyran [7] showed when, in a quasi metric space, a generalized inexact proximal algorithm, equipped with a generalized perturbation term $\Gamma\left[q(x,y)\right]$, and defined, each step by, i) a sufficient descent condition and, ii) a stopping rule, converges to a critical point. Then, they have shown that the speed of convergence and convergence in finite time depends of the curvature of the perturbation term and of the Kurdyka-Lojasiewicz property associated to the objective function. A striking and new application has been given. It concerns the impact of the famous "loss aversion effect" (Nobel Prize [8, 9]) on the speed of convergence of the generalized inexact proximal algorithm. However, in the context of the "Variational rationality approach", which considers, as central dynamical concepts, worthwhile stay and change processes, these important results in Applied Mathematics are not enough, from the viewpoint of our applications to Behavioral Sciences, unless we can show that this critical point is a variational trap (strong or weak) where the agent will prefer to stay than to move, because his motivation to change is strictly or weakly lower than his resistance to change. This section presents, under the conditions of [7, Theorem 3.1], a worthwhile stay and change process which converges to a critical point of f which is a weak trap (compare, below, with the definition of a strong global trap). Then, start with the general definition of a weak global trap instead of a strong one ([1, 2]).

Definition 4.1. *Let $x \in X$ be a given action and $\xi > 0$ be a satisficing rate of change choosen by the agent. Let $W_{\xi}(x) := \{y \in X, M(x, y) \geq \xi R(x, y)\}$ be his worthwhile to change set, starting from $x \in X$. Then, starting from $x^* \in X$ with a given satisficing worthwhile to change rate $\xi^* > 0$, a strong variational trap $x^* \in X$ is such that motivation to change is stricly lower than resistance to change, $M(x^*, y) < \xi_* R(x^*, y)$ for all $y \neq x^* \in X$. A weak variational trap is such that $M(x^*, y) \leq \xi_* R(x^*, y)$, for all $y \in X$. This defines the stationary side of a trap. The variational aspect comes from being the end of a worthwhile to change process, starting from an initial given point.*

Remark 4.1.

a) Notice that a strong global trap is such that $W_{\xi_*}(x^*) = \{x^*\}$ and a weak global trap is such that $W_{\xi_*}(x^*) = \{y \in X,\ M(x^*, y) = \xi_* R(x^*, y)\}$. At a strong (weak) global trap, the agent strictly (weakly) prefers to stay than to move. Then, when a process of worthwhile stays and changes converges to a strong variational trap, this variational formulation defines, starting from an initial point, a variational trap as the end point of a path of worthwhile changes, worthwhile to approach, but not worthwhile to *leave. This because, starting from there, there is no way to do any other worthwhile change, except repetitions.*

b) *Assuming that* $\{\lambda_k\}$ *converges to* λ_∞*, our sufficient condition proposes an algorithm which, following a succession of worthwhile changes* $x^{k+1} \in W_{\lambda_k}(x^k), k \in \mathbb{N}$*, converges to a weak global trap* x^* *such that* $W_{\lambda_\infty}(x^*) = \{y \in X : \ M(x^*, y) = \lambda_\infty R(x^*, y)\}$*. Since the agent is free to choose all his satisficing worthwhile to change rates* λ_k *in an adaptive way, this will show that the agent, choosing at the limit point* x^* *a satisicing worthwhile to change rate* $\lambda_* > \lambda_\infty$*, ends in a strong global trap* x^**, because* $M(x^*, y) = \lambda_\infty R(x^*, y) < \lambda_* R(x^*, y)$*, for all* $y \in X$*.*

c) *As observed in Section 2, in the specific context of this paper, we have*

$$M(x, y) = U\left[A(x, y)\right] = f(x) - f(y)$$

$$R(x, y) = D\left[C(x, y)\right], \quad \Gamma\left[q(x, y)\right] = \lambda U^{-1}\left[D[q(x, y)]\right], \xi = 1.$$

Then, in our present paper, a strong (resp. weak) variational trap is such that $f(x^*) - f(y) < \lambda\Gamma\left[q(x^*, y)\right]$*, for all* $y \neq x^*$ *(resp.* $f(x^*) - f(y) \leq \lambda\Gamma\left[q(x^*, y)\right]$*, for all* $y \in X$*).*

4.2 Some Definitions from Subdifferential Calculus

In this section some elements concerning the subdifferential calculus are recalled; see, for instance, [22, 23]. Assume that $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is a proper lower semicontinuous function. The domain of f, which we denote by $\mathrm{dom} f$, is the subset of $\mathbb{R}^n$ on which f is finite-valued. Since f is proper, then $\mathrm{dom} f \neq \emptyset$.

Definition 4.2.

i) The Fréchet subdifferential of f at $x \in \mathbb{R}^n$, denoted by $\hat{\partial} f(x)$, is the set given by:

$$\hat{\partial} f(x) := \begin{cases} \{x^* \in \mathbb{R}^n : \liminf_{y \to x; y \neq x} \frac{1}{\|x - y\|}(f(y) - f(x) - \langle x^*, y - x\rangle) \geq 0\}, & if\ x \in domf, \\ \emptyset, & if\ x \notin domf. \end{cases}$$

ii) The limiting Fréchet subdifferential (or simply subdifferential) of f at $x \in \mathbb{R}^n$, denoted by $\partial f(x)$, is the set given by:

$$\partial f(x) := \begin{cases} \{x^* \in \mathbb{R}^n | \exists x_n \to x, f(x_n) \to f(x), x_n^* \in \hat{\partial} f(x_n); x_n^* \to x^*\}, & if\ x \in domf. \\ \emptyset, & if\ x \notin domf. \end{cases}$$

Throughout the paper we consider the subdifferential ∂f since it satisfies a closedness property important in our convergence analysis, as well as in any limiting processes used in an algorithmic context.

A necessary condition for a given point $x \in \mathbb{R}^n$ to be a minimizer of f is

$$0 \in \partial f(x). \tag{8}$$

It is known that, unless f is convex, (8) is not a sufficient condition. The domain of ∂f, which we denote by dom ∂f, is the subset of $\mathbb{R}^n$ on which ∂f is a nonempty set. In the remainder, a point that satisfies (8) is called limiting-critical or simply critical point.

4.3 The Algorithm

In [3] the authors examined the "local epsilon inexact proximal" algorithm,

$$f(x^{k+1}) + \lambda_k d(x^k, x^{k+1}) \leq f(y) + \lambda_k d(x^k, y) + \varepsilon_k, \qquad y \in E(x^k, r_k) \subset X,$$

where, i) d is a distance, ii) $E(x^k, r_{k+1}) \subset X$ is a variable choice set (a moving ball), for each current period $k+1$. Following [3] we consider the so called "global epsilon inexact proximal" algorithm as follows: starting from the current position x^k, let us define the next iterate x^{k+1} as follows:

$$f(x^{k+1}) + \lambda_k \Gamma[q(x^k, x^{k+1})] \leq f(y) + \lambda_k \Gamma[q(x^k, y)] + \varepsilon_k, \qquad y \in X, \tag{9}$$

where $\{\lambda_k\}, \{\varepsilon_k\}$ are given sequences of nonnegative real numbers, and q is a quasi distance. In the particular case where the generalized perturbation term $\Gamma[q(x,y)] = q(x,y)^2$ and $q(x,y) = d(x,y)$ is a distance, instead of a quasi distance, our "global epsilon inexact proximal" algorithm coincides with the case considered by Zaslavski [24].

Assumption 4.1. *There exist $\beta_1, \beta_2 \in \mathbb{R}_{++}$ such that: $\beta_1\|x-y\| \leq q(x,y) \leq \beta_2\|x-y\|$, $x, y \in \mathbb{R}^n$.*

This is the case in our knowledge management example. For an other explicit example where inconvenients to change are a quasi-distance satisfying Assumption 4.1, see [12].

Next, we recall the inexact version of the proximal point method introduce in [7].

Algorithm 4.1. *Take $x^0 \in domf$, $0 < \bar{\lambda} \leq \tilde{\lambda} < +\infty$, $\sigma \in [0,1[$ and $b > 0$. For each $k = 0, 1, \ldots$, choose $\lambda_k \in [\bar{\lambda}, \tilde{\lambda}]$ and find $(x^{k+1}, w^{k+1}, v^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ such that:*

$$f(x^k) - f(x^{k+1}) \geq \lambda_k(1-\sigma)\Gamma[q(x^k, x^{k+1})], \tag{10}$$

$$w^{k+1} \in \partial f(x^{k+1}), \quad v^{k+1} \in \partial q(x^k, \cdot)(x^{k+1}), \tag{11}$$

$$\|w^{k+1}\| \leq b\Gamma'[q(x^k, x^{k+1})]\|v^{k+1}\|, \tag{12}$$

The first condition is a sufficient descent condition. It is a (proximal-like) worthwhile to change condition $x^{k+1} \in W_{\xi_{k+1}}(x_k)$, where the proximal perturbation term defines the relative resistance to change function. This condition tells us that it is worthwhile to change from x^k to x^{k+1}, rather than to stay at x^k. In this case, advantages to change from x^k to x^{k+1}, $A(x^k, x^{k+1}) = f(x^k) - f(x^{k+1})$ are, each period, higher than some adaptive proportion $\xi_{k+1} = \lambda_k(1-\sigma)$ of the relative disutility of inconvenients to change rather than to stay $\Gamma[q(x^k, x^{k+1})] = U^{-1}\left[D\left[I(x^k, x^{k+1})\right]\right]$, where, i) inconvenients to change rather than to stay are $I(x^k, x^{k+1}) = C(x^k, x^{k+1}) - C(x^k, x^k) = q(x^k, x^{k+1})$, ii) costs to be able to change from x^k to x^{k+1} are $C(x^k, x^{k+1}) = q(x^k, x^{k+1})$, while, iii) costs to be able to stay $C(x^k, x^k) = q(x^k, x^k) = 0$ are zero as quasi *distances. The second conditions defines subgradients of the objective and costs to be able to change functions. The third condition is a stopping rule which says, each period, when the agent prefers to do not make a new marginal change, because it is not worthwhile to do it, this period; see Section 3.2 on marginally worthwhile changes.*

Remark 4.2. *As pointed out by the authors, Algorithm 4.1 retrieves the inexact algorithm proposed in [11, Algorithm 2] in the particular case* $\Gamma[q] = q^2/2$, $q(x,y) = \|x-y\|$ *and* $1-\sigma = \theta$. *Moreover, Algorithm 4.1 is an habituation/routinization process and any sequence generated from it is a path of worthwhile changes with parameter* $\lambda_k(1-\sigma)$ *such that, at each step, it is marginally worthwhile to stop. The variational stopping rule condition raises the following question: when, marginally, a change stops to be worthwhile? This strongly depends on the shapes of the utility and desutility functions.*

Comparing Algorithm 4.1 with the iterative process (9), we observe the following:

i) on one side, the iterative process (9) is much more specific than our Algorithm 4.1. Indeed the weak "worthwhile to change" condition (10) is replaced by the much stronger condition (9).

ii) on the other side, the iterative process (9) does not impose the "not worthwhile marginal change condition" (12) as the Algorithm 4.1 does.

Next we propose a new inexact proximal algorithm, combining a particular instance of (9) with the stopping rule (12).

Algorithm 4.2. *Take* $x^0 \in dom f$, $0 < \bar{\lambda} \leq \tilde{\lambda} < +\infty$, $\sigma \in [0,1[$ *and* $b > 0$. *For each* $k = 0, 1, \ldots$, *choose* $\lambda_k \in [\bar{\lambda}, \tilde{\lambda}]$ *and find* $(x^{k+1}, w^{k+1}, v^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ *such that:*

$$f(y) - f(x^{k+1}) \geq \lambda_k \left[(1-\sigma)\Gamma[q(x^k, x^{k+1})] - \Gamma[q(x^k, y)]\right], \qquad y \in X, \tag{13}$$

$$w^{k+1} \in \partial f(x^{k+1}), \quad v^{k+1} \in \partial q(x^k, \cdot)(x^{k+1}), \tag{14}$$

$$\|w^{k+1}\| \leq b\Gamma'[q(x^k, x^{k+1})]\|v^{k+1}\|. \tag{15}$$

Remark 4.3. *This new inexact proximal algorithm imposes a stronger worthwhile to change condition than Algorithm 4.1, because it must be verified, each period, for each $y \in X$. Setting $y = x^k$ gives the last worthwhile to change condition. The other two conditions remain unchanged. Note that the exact proximal algorithm* (3) *is a specific case of our new algorithm (it holds by taking $\sigma = 0$). The new inexact worthwhile to change condition is $P_{\lambda_k}(x^k, x^{k+1}) \leq P_{\lambda_k}(x^k, y) + \lambda_k \sigma \Gamma\left[q(x^k, x^{k+1})\right]$, for all $y \in X$.*

As in [10, 12, 11, 7], our main convergence result is restricted to functions that satisfy the so-called Kurdyka-Lojasiewicz inequality; see, for instance, [25, 26, 27, 28]. Next formal definition of the Kurdyka-Lojasiewicz inequality can be finding in [28], where it is also possible to find several examples and a good discussion over important classes of functions which satisfy the mentioned inequality.

Definition 4.3. *A proper lower semicontinuous function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is said to have the Kurdyka-Lojasiewicz property at $\bar{x} \in dom\ \partial f$ if there exists $\eta \in]0, +\infty]$, a neighborhood U of $\bar{x}$ and a continuous concave function $\varphi : [0, \eta[\to \mathbb{R}_+$ such that:*

$$\varphi(0) = 0, \quad \varphi \in C^1(0, \eta), \quad \varphi'(s) > 0, \quad s \in]0, \eta[; \tag{16}$$

$$\varphi'(f(x) - f(\bar{x}))dist(0, \partial f(x)) \geq 1, \quad x \in U \cap [f(\bar{x}) < f < f(\bar{x}) + \eta], \tag{17}$$

- $dist(0, \partial f(x)) := inf\{\|v\| : v \in \partial f(x)\}$,
- $[\eta_1 < f < \eta_2] := \{x \in M : \eta_1 < f(x) < \eta_2\}, \quad \eta_1 < \eta_2.$

In what follows, we assume that f a is bounded from below, continuous on dom$f$ and KL function, i.e., a function which satisfies the Kurdyka-Lojasiewicz inequality at each point of dom∂f.

Theorem 4.1. *Assume that $\{x^k\}$ is bounded sequence generated from Algorithm 4.2, $\tilde{x}$ is an accumulation point of $\{x^k\}$ and Assumption 4.1 holds. Let $U \subset \mathbb{R}^n$ be a neighborhood of $\tilde{x}$, $\eta \in]0, +\infty]$ and $\varphi : [0, \eta[\to \mathbb{R}_+$*

a continuous concave function such that (16) *and* (17) *hold. If* $\delta \in (0, \bar{q})$ (*see condition* (5)) *and* $r \in]0,1[$ *are fixed constants,* $B(\tilde{x}, \delta/\beta_1) \subset U$*,* $a := \bar{\lambda}(1-\sigma)$ *and* $M := \frac{Lb}{a}$*, then the whole sequence* $\{x^k\}$ *converges to a critical point* x^* *of* f *which is a strong global trap, relative to the worthwile to change set* $W_{\lambda_*}(x^*)$*, for any choice of the final satisficing rate* $\lambda_* > \lambda_\infty$*.*

Proof. The first part of the theorem follows immediately from [7, Theorem 3.1] because any sequence, generated from Algorithm 4.2, satisfies the conditions (10) and (12) of Algorithm 4.1. Let x^* be the limit point of the sequence $\{x^k\}$. Given that the sequence $\{\lambda_k\} \subset [\bar{\lambda}, \tilde{\lambda}]$ (is bounded), $0 < \bar{\lambda} \leq \tilde{\lambda} < +\infty$, taking a subsequence, if necessary, we can assume that λ_k converges to a certain $\lambda_\infty \in]0, +\infty[$. For the second part, note that $\{f(x^k)\}$ is a non increasing sequence and $x^* \in \text{dom} f$. Now, given that $q(\cdot, y)$ is continuous for each $y \in X$ (see [12]), Γ is continuous and f is continuous on $\text{dom} f$, taking the limit in (13) as k goes to infinity and assuming that λ_k converges to a certain $\lambda_\infty \in]0, +\infty[$, we get:

$$f(x^*) \leq f(y) + \lambda_* \Gamma[q(x^*, y)], \qquad y \in X.$$

Therefore, the desired result follows from Remark 4.1. □

5 Application to the Formation and Break of Habits/Routines

To save space, this section represents a really very short summary of the "available to read? long preprint of our present paper, named "Some comparisons between the Variational rationality, Habitual domain and DMCS approaches"; see [13]. The *Variational rationality (VR) approach* (see [1, 2]) focus attention on interdisciplinary stability and change dynamics (habits and routines, creation and innovation, exploration and exploitation,. . .), and the self regulation problem, seen as a stop and go course pursuit between feasible means and desirable ends mixing, in alternation, discrepancy production (goal setting, goal revision), discrepancy reduction (goal striving, goal pursuit) and goal disengagement. It rests on two main concepts, worthwhile temporary stays and changes, variational traps and nine principles. This (VR) approach allows to recover the main mathematical variational principles, and in turn, it benefits from almost all variational algorithms for procedural applications, which all, use some of the main variational rationality principles. The *Habitual domain (HD) theory and (DMCS) approach* (see Yu and Chen [29], for a nice presentation) and the (DMOCS) Decision making and optimization problems in changeable spaces (see Larbani and Yu [15]) refer to three stability and change problems. They are i) stability issues, using a system of differential equations, a variant of the famous pattern formation Cohen-Grossberg model (see Cohen and Grossberg [30]), ii) expansion of an initial competency set to be able to solve a given problem, which requires to acquire a new given competency set, using mathematical programming methods and graphs, iii) DMCS optimization and game problems, using Markov chains, with applications to innovation cover-discover problems (see Yu and Larbani [31] and Larbani and Yu [32, 33]).

1) Habit/routine formation and break problems

The (VR) approach sees habit/routine formation and break as a balance between motivation and resistance to change, when agents use worthwhile changes; Permanent habits refer to variational traps, as the end of a succession of worthwhile changes. These findings fit well with Psychology and Management theories of habits and routines formation and break. For example, in Psychology, habits form by repetitions, in a rather stable context, which trigger their repetition. They become gradually more and more automatized behaviors which are intentional (goal directed), more or less conscious and controllable, and economize cognitive resources. In this context agents are bounded rational. The (HD) approach modelizes habit formation as a balance, each step, between excitation and inhibition forces, which determine, each time, the variation of the allocation of attention and effort (propensities), using, as said before, a variant of the Cohen-Grossberg system of differential equations (see [30]).

In the limit, the allocation of attentions and efforts (propensities) converge to an allocation which represents a stable habitual domain. The (DMOCS) approach of games defines limit profiles of mind sets as absorbing states of a Markow chain. In these two contexts agents are bounded rational.

2) Inexact proximal algorithms as repeated satisficing problems with changeable decision sets

We have shown in Section 3 of this paper that our inexact proximal algorithms refer to variable and changeable decision sets, payoffs, goals (satisficing worthwhile changes) and preference processes. This is the case because worthwhile to change sets are changeable decision sets which change, each period, with experience and the choice, each period, of the satisficing worthwhile to change ratio; see the whole Section 2 and the dynamic of worthwhile (hence satisficing) temporary stays and changes. More precisely Inexact generalized proximal algorithms are specific instance of VR worthwhile stay and change adaptive dynamics. They consider non transitive variable worthwhile to change preferences, see Section 2, which are reference dependent preferences with variable reference points (variable experience dependent utility/disutility functions, variable payoff functions, variable and non linear resistance to change functions via, in each case, of the introduction of the separable and variable lambda term). They can use costs to be able to change which do not satisfy the triangular inequality ([34] and several other references within [13]). They are inexact and procedural algorithms on quasi metric spaces. Then, they modelize bounded rational and more or less myopic agents who bracket difficult decisions in several steps, contrary to exact proximal algorithms which modelize a repeated optimization problem, and which are not the topic of our paper. They generalize the Simon satisficing principle to a dynamical context (repeated and adaptive satisficing). They are anchored to optimization processes as benchmark cases. They deal with changeable spaces as changeable worthwhile to change sets and also changeable exploration sets (see [3, 34]) and consider convergence in finite time as a central topic (see [7, 35]). They include psychological aspects (like motivations, cognitions and inertia) and can easily include emotional aspects.

We point out that the assumptions of inexact proximal algorithms explicit in behavioral terms have been given in Section 2 of this paper (see the example) and after each proximal algorithm.

6 Conclusion

In this paper, following [7], and using the recent variational approach presented in [1, 2], we have proposed a generalized "epsilon inexact proximal" algorithm that converges to a critical point which is also a variational trap. In Mathematics, our paper helps to show how the literature on proximal algorithms can be divided in two parts: the case of strong and weak relative resistance to change. In this paper we have considered the most difficult situation, the weak case. In Behavioral Sciences, our paper offers a dynamic model for habituation/routinization processes, and gives a striking and new result on the impact of the famous "loss aversion" index (see the Nobel Prize [8, 9]) on the speed of convergence of such processes. Given editorial constraints (lack of space in the present paper), this important result appears in the first paper ([7]). In [13] the authors compare our VR variational rationality approach of inexact proximal algorithms to the HD habitual domain theory and DMCS approach ([14, 15]). Future research will consider the multiobjective case.

References

1. Soubeyran, A.: Variational rationality, a theory of individual stability and change: worthwhile and ambidextry behaviors. Pre-print. GREQAM, Aix Marseillle University, (2009).

2. Soubeyran, A.: Variational rationality and the "unsatisfied man": routines and the course pursuit between aspirations, capabilities and beliefs. Preprint. GREQAM, Aix Marseillle University (2010).

3. Attouch.H, Soubeyran, A.: Local Search Proximal Algorithms as Decision Dynamics with Costs to Move. Set-Valued Var. Anal. **19**(1), 157–177 (2011)

4. Flores-Bazán.F., Luc. D, Soubeyran, A.: Maximal Elements Under Reference-Dependent Preferences with Applications to Behavioral Traps and Games. J Optim Theory Appl. **155**, 883–901 (2012)

5. Attouch, H., Redont, P., Soubeyran, A.: A new class of alternating proximal minimization algorithms with costs-to-move, SIAM J. Optimiz. **18**, 1061-1081 (2007)

6. Cruz Neto, J. X., Oliveira, P. R., Soares Jr, P. A., Soubeyran, A.: Learning how to play Nash, potential games and alternating minimization method for structured nonconvex problems on Riemannian manifolds. To appear in J. Convex Anal. (2013)

7. Bento, G. C., Soubeyran, A.: A Generalized Inexact Proximal Point Method for Nonsmooth Functions that Satisfies Kurdyka Lojasiewicz Inequality. Submitted (2014)

8. Kahneman, D., Tversky, A.: Prospect theory: An analysis of decision under risk. Econometrica **47**(2), 263-291 (1979)

9. Tversky, A., Kahneman, D.: Loss aversion in riskless choice: a reference dependent model. Q. J. Econ. **106**(4), 1039-1061 (1991)

10. Attouch, H., Bolte, J.: On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. Math. Programming, Ser. B. **116**(1-2), 5-16 (2009)

11. Attouch, H., Bolte, J., Svaiter, B. F.: Convergence of descent methods for semi-algebric and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods. Math. Programming, Ser. A. **137**, 91-129 (2012)

12. Moreno, F. G., Oliveira, P. R., Soubeyran, A.: A Proximal Algorithm with Quasi Distance. Application to Habit's Formation. Optimization **61**(12), 1383-1403 (2011)

13. Bento, G. C., Soubeyran, A.: Inexact proximal algorithms in models of Behavioral Sciences. Arxiv, arXiv:1403.7032 [math.OC] (2014)

14. P. L. Yu, Forming Winning Strategies, An Integrated Theory of Habitual Domains, Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo (1990)

15. Larbani, M., Yu, P. L.: Decision Making and Optimization in Changeable Spaces, a New Paradigm. J. Optim. Theory Appl., **155**(3), 727-761 (2012)

16. Simon. H.: A behavioral model of rational choice. Q. J. Econ. **69**, 99-118 (1955)

17. Van Long, N., Soubeyran, A., Soubeyran, R.: Knowledge Accumulation within an Organization. To appear in Int. Econ. Rev. (2014)

18. Kremer, M.: The O-ring theory of economic development. Quart. J. Econ. **108**, 551?575 (1993)

19. Rockafellar, R. T.: Monotone operators and the proximal point algorithm. SIAM J. Control. Optim. **14**, 877-898 (1976)

20. Moreau, J.: Proximité et dualité dans un espace hilbertien. (French) Bull. Soc. Math. **93**, 273-299 (1965)

21. Martinet, B.: Régularisation, d'inéquations variationelles par approximations successives. (French) Rev. Française Informat. Recherche Opérationnelle **4**, Ser. R-3, 154-158 (1970)

22. Rockafellar, R. T., Wets, R.: Variational Analysis, **317** of Grundlehren der Mathematischen Wissenschafte. Springer, 1998 (3rd printing 2009).

23. Mordukhovich, B.: Variational Analysis and Generalized Differentiation I: Basic Theory. Springer, (Grundlehren der mathematischen Wissenschaften) (2010)

24. Zaslavski, A. J.: Inexact Proximal Point Methods in Metric Spaces. Set-Valued Anal. **19**, 589-608 (2011)

25. Lojasiewicz, S.: Une propriété topologique des sous-ensembles analytiques réels. Les Équations aux Dérivées Partielles, Éditions du centre National de la Recherche Scientifique, 87-89 (1963)

26. Kurdyka, K.: On gradients of functions definable in o-minimal structures. Ann. Inst. Fourier, **48**, 769-783 (1998)

27. Bolte, J., Daniilidis, J. A., Lewis, A.: The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**(4), 1205-1223 (2006)

28. Attouch, H., Redont, P., Bolte, J., Soubeyran, A.: Proximal Alternating Minimization and Projection Methods for Nonconvex Problems. An Approach Based on the Kurdyka-Lojasiewicz Inequality. Math. Oper. Res. **35**(2), 438-457 (2010)

29. Yu, P. L., Chen,Y. C.: Dynamic MCDM, habitual domains and competence set analysis for effective decision making in changeable spaces. Chaper 1. In Trends in Multiple Criteria Decision Analysis. Matthias Ehrgott Jose Rui Figueira Salvatore Greco. Springer (2010)

30. Cohen, M., Grossberg, S.: Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. IEEE T. Syst. Man. CYB, **13**, 815-826 (1983)

31. Yu, P. L., Larbani M.: Two-person second-order games, Part 1: formulation and transition anatomy. J. Optim. Theory Appl. **141**(3), 619-639 (2009)

32. Larbani, M., Yu, P. L.: Two-person second-order games, Part II: restructuring operations to reach a win-win profile. J. Optim. Theory Appl., **141**, 641-659 (2009)

33. Larbani, M., Yu, P.L.: n-Person second-order games: A paradigm shift in game theory. J. Optim. Theory Appl., **149**(3), 447–473 (2011)

34. Bento, G. C., Cruz Neto, J. X., Soubeyran, A.: A Proximal Point-Type Method for Multicriteria Optimization. In revision, Set-Valued Var. Anal. (2014)

35. Bento, G. C., Cruz Neto, J. X., Soares Jr., P.A., Soubeyran, A.: Behavioral Traps and the Equilibrium Problem on Hadamard Manifolds. Arxiv, arXiv:1307.7200v2 [math.OC] (2013)