A Dynamical System Based Heuristic for a Class of Constrained Semi-Assignment Problem

Michael Burkard, Swiss Federal Institute of Technology, Zürich

Maurice Cochand, Swiss Federal Institute of Technology, Zürich

Arlette Gaillard, Swiss Federal Institute of Technology, Zürich

Summary: We consider Constrained Semi-Assignment Problems, where one has to optimally assign discrete values to decision variables, and where furthermore logical constraints restrict the set of feasible assignments. We present a heuristic approach based on a continuous relaxation. More precisely, we consider a discrete dynamical system evolving in the interior of a polytope whose extremal points correspond to all possible (both feasible and infeasible) assignments. The constructed dynamical system combines the effects of two dynamics: The first one, being of gradient-type, has the task of attracting the system towards assignments with high objective value, the second dynamic should give a rejecting power to forbidden assignments. Local properties of the system and numerical experiments are discussed.

1. Introduction

The constrained maximization problem considered in this paper generalizes the well known pseudoboolean optimization and satisfiability (SAT) problems. Instead of boolean variables we consider decision variables having a finite set of possible values. As in SAT, a set of logical clauses is given and those assignments satisfying all given clauses constitute the set of feasible solutions of our problem. Furthermore, each assignment is given an objective value by means of a set of weighted clauses: the value of an assignment is the sum of the weights of the clauses it satisfies. The constrained semi-assignment problem (C-SAP) consists in finding a feasible assignment of maximal value. For the formal definition let (i, r) stand for the predicate: value r is assigned to variable x_i.

Constrained semi-assignment problem (C-SAP): Given

- a set of decision variables $x_i, i \in N := \{1, \dots, n\}$ with possible values in $K := \{1, 2, \dots, k\}$,

- a set of clauses $\mathcal{R}$, where $R \in \mathcal{R}$ is of the form

$$R = \neg\left((i_1, r_1) \wedge (i_2, r_2) \wedge \dots \wedge (i_u, r_u)\right)$$

  for some $2 \leq u \leq n, i_j \in N, r_j \in K$, for all $j \leq u$. For convenience, R will be used to denote the set $R = \{(i_1, r_1), (i_2, r_2), \dots, (i_u, r_u)\}$ as well.

- a set of weighted clauses $\mathcal{T}$ with weights $w_T > 0, T \in \mathcal{T}$, where T is of the form

$$T = (i_1, r_1) \wedge (i_2, r_2) \wedge \dots \wedge (i_u, r_u)$$

  for some $u \leq n, i_j \in N, r_j \in K$, for all $j \leq u$. Again, T is identified with the set $T = \{(i_1, r_1), (i_2, r_2), \dots, (i_u, r_u)\}$.

Find an assignment of values in K to the variables satisfying all clauses of $\mathcal{R}$ and maximizing

$$\sum\{w_T | T \in \mathcal{T}, T \text{ satisfied by the assignement}\}.$$

Note that for $k = 2$ and $\mathcal{T} = \emptyset$ C-SAP reduces to the classical satisfyability problem; for k = 2 and $\mathcal{R} = \emptyset$ we have a pseudo-boolean optimization (see formulation (1)-(3)). Many other combinatorial problems can be stated in this form, like max k-cut, k-coloring and max clique. C-SAP contains therefore many NP-hard problems, including some for which even finding a poor approximate solution is NP-hard. Moreover, we believe that C-SAP is an adequate formalism for many interesting applications, like for example the label placement problem where a finite set of possible positions is defined for each label and additionaly to the natural non intersecting clauses for the labels, a valuation is introduced according to aesthetical criteria or to the ambiguities resulting from the placements.

A common and often successful approach for such problems are local search (LS) heuristics, like Simulated Annealing or Tabu Search. However, due to their inherent myopy, such heuristics can hardly deal with a C-SAP of the following paradigmatic type:

Let $D \subseteq K^N$ be given and $y \in D$ be a unique isolated global optimum with objective value $z(y) = M \gg 0$. Moreover, let $z(x) = 0$ for all $x \in D\backslash\{y\}$ (C-SAP with $\mathcal{T} := \{T := \{(1, y_1), (2, y_2), \ldots, (n, y_n)\}\}$ and $\mathcal{R} := \left\{\{(1, z_1), \ldots, (n, z_n)\} | z \in K^N \setminus D\right\}$). Our goal is to present a heuristic with some ability to 'orient itself' even in such problems and which therefore offers an alternative approach for instances, where LS algorithms have problems of orientation. In fact, we will show that if M is big enough in the above instance our algorithm will find the global optimum (see Theorem 2.1).

Our approach is to embed the set of feasible solutions in the continuous space and to develop a heuristic based on a discrete time dynamical system. The concept of our approach can be formulated as follows: We represent the set of (a priori) possible assignments $x \in K^N$ as points in

$$\Delta^I := \{p \in \{0,1\}^{N \times K} | \sum_{r=1}^{k} p_{ir} = 1 \text{ for all } i \in N\},$$

where $p \in \Delta^I$ corresponds to the assignment $x_i = r$ iff $p_{ir} = 1$. Then a C-SAP problem (with sets $\mathcal{R}$ and $\mathcal{T}$) can be written as

$$\max \quad z(p) = \sum_{T \in \mathcal{T}} \left(w_T \prod_{(i,r) \in T} p_{ir} \right) \tag{1}$$

$$\text{s.t.} \quad \prod_{(i,r) \in R} p_{ir} = 0 \qquad \text{for all } R \in \mathcal{R} \tag{2}$$

$$p \in \Delta^I \tag{3}$$

Let now $\Delta := \text{Conv}(\Delta^I)$ and Δ^0 be its interior. Our approach consists in defining a mapping $F : \Delta^0 \to \Delta^0$ and to use the discrete time dynamical system resulting by iteration of F for our algorithm which runs essentially as follows: Choose a point $p^0 \in \Delta^0$ and compute the sequence $p^i := F(p^{i-1}), i = 1, 2, ..., t$ for some t. Return as solution $p \in \Delta^I$ the solution p "closest" to p^t.

The underlying philosophy is that optimal solutions of the combinatorial problem should correspond to attractors of the dynamical systems. The definition of F is of course central and we would be already happy to have the sequence p^i converging to one of those points in Δ^I corresponding to a 'good' feasible solution for a reasonably large proportion of starting points. The list of desirable properties of such a dynamic includes furthermore the instability of non interpretable fixed points in Δ^0, the escape from some neighborhood of non feasible assignments (in particular of those having a feasible neighbor), no matter the value of the assignment.

The application of dynamical systems in combinatorial optimization is definitely not a popular approach, however the works of [2] for Maxclique, [4] for Maxsat, and [10] for combinatorial problems reducible to proper assignment problems (i.e assignment in the sense of permutation) show that it has some potential to be explored (see also the book of [6]).

In the first part of this paper we address the design of a mapping F for C-SAP and discuss some of its basic properties. In the second part we shall report on some numerical experiments.

2. Algorithm

The Algorithm consists in defining an adequate operator $F : \Delta^0 \to \Delta^0$ for a given C-SAP problem and to generate a sequence of points $p^0, p^1 = F(p^0), \ldots, p^t$ for some t and to round p^t to a solution p in Δ^I. For F we consider mappings defined by

$$F(p)_{ir} := \frac{p_{ir}\xi_{ir}(p)}{N_i} \text{ with } N_i := \sum_r p_{ir}\xi_{ir}(p) \tag{4}$$

where the so-called fitness functions $\xi_{ir}(p) > 0$ are given. As we shall consider different possibilities for ξ_{ir}, we shall use the notation $F[\xi]$ for the mapping defined by (4). In the setting of multi-population dynamic where n populations are considered whose members can adopt one of k possible strategies, this kind of dynamic is referred to as "adjusted replicator dynamic" and describes the evolution of the proportion of individuals in population i adopting strategy r (see [9]). This evolution equation expresses the fact that the relative increase $(F(p)_{ir} - p_{ir})/p_{ir}$ of p_{ir} equals the relative excess fitness $\frac{\xi_{ir}(p)-N_i}{N_i}$ of subpopulation (i, r) over the average fitness of population i.

We shall combine in our fitness function the effects of two dynamics: The first, being of gradient-type, has the task of attracting the system towards assignments with high objective values, the second dynamic should care for making logically forbidden assignments repellors. Ideally, in the combined dynamic, the effect of the first dynamic should be stronger in the neighborhood of admissible assignments, and "vice versa". For the gradient-type dynamic we choose $F[\Gamma^\alpha]$ with

$$\Gamma_{ir} \;:=\; \partial z/\partial p_{ir} = \sum_{T:(i,r)\in T} w_T \prod_{(j,s)\in T\setminus(i,r)} p_{js}$$

where z is the objective function (1). For the special case $\alpha = 1$, this dynamic has been studied in [1] and it has been shown that in this case $z(F[\Gamma](p)) \geq z(p), p \in \Delta^0$, with equality for fixed points only. (In fact this result remains true, if $\alpha \leq 1$, see Theorem 2.4.)

For the second dynamic, similarly to [4] we choose $F[\Theta^\beta]$, where

$$\Theta_{ir} \;:=\; \prod_{R:(i,r)\in R} (1 - \prod_{(j,s)\in R\setminus(i,r)} p_{js}).$$

Note that if p is near $\bar{p} \in \Delta^I$ and $\bar{p}$ violates some constraint R, then for $(i, r) \in R$ the fitness $\Theta_{ir}(p)$ is as desired very small, what should contribute to make p_{ir} smaller.

The combined dynamic F is then defined as

$$F(p)_{ir} := F[\Gamma^\alpha \Theta^\beta](p)_{ir} = \frac{p_{ir}\Gamma_{ir}(p)^\alpha \cdot \Theta_{ir}(p)^\beta}{\sum_{s=1}^{k} p_{is}\Gamma_{is}(p)^\alpha \cdot \Theta_{is}(p)^\beta} \tag{5}$$

In Figure 1 the effects of the three dynamics are shown for the C-SAP with $n = k = 2$

$$\begin{aligned} \max \quad & z(p) = 4p_{11}p_{21} + 5p_{12}p_{22} + 2p_{12}p_{21} + 3p_{11}p_{22} \\ \text{s.t.} \quad & p_{11}p_{21} = 0, \quad p_{12}p_{22} = 0, \quad p \in \Delta^I \end{aligned}$$

In the first picture we show the trajectories for the gradient-type dynamic $F[\Gamma^{0.5}]$ used to maximize $z(p)$ (without satisfyability constraints). The dynamic $F[\Theta]$ adequate to "flow away" from the forbidden assignments (without optimization) is represented in the second picture and the trajectories of a combined dynamic $F[\Gamma^2\Theta^{0.5}]$ are shown in the last picture.

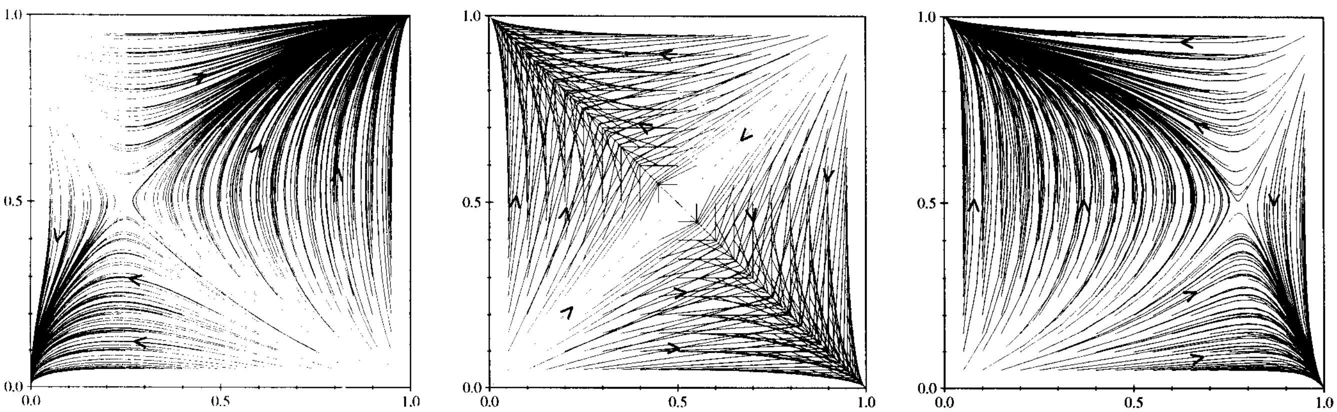

Figure 1: Trajectories of $F[\Gamma^{0.5}]$, $F[\Theta]$ and of the combined dynamic $F[\Gamma^2\Theta^{0.5}]$

Note that $F[\Theta]$ cannot be extended to the boundary of Δ for some denominator may vanish at infeasible assignments. For theoretical reasons (see Theorem 2.2) we shall restrict the domain of all the operators involving Θ to

$$\Delta^\varepsilon := \{p \in [\varepsilon,\ 1 - (k-1)\varepsilon]^{N \times K} | \sum_{r=1}^{k} p_{ir} = 1 \text{ for all } i \in N\}.$$

If not discussing the current value of ε we consider ε as fixed. The range of our operator must be Δ^ε as well and we define $\Psi : \Delta \to \Delta^\varepsilon$ as follows: If $p_{ir} < \varepsilon$ then $\Psi(p)_{ir} := \varepsilon$, $\Psi(p_{ir(i)}) := \min\{p_{ir(i)}, 1 - (k-1)\varepsilon\}$, where $p_{ir(i)} > p_{is}$, $s < r(i)$ and $p_{ir(i)} \geq p_{is}$, $s > r(i)$ and finally for the remaining components $\Psi(p)_{ir} := \varepsilon + \mu(p_{ir} - \varepsilon)$, with μ chosen appropriate to fulfill $\Psi(p) \in \Delta^\varepsilon$. We denote

$$\bar{F}[\xi] := \Psi \circ F[\xi] \tag{6}$$

and define $Round : \Delta \to \Delta^I$ by $Round(p)_{ir} = 1$ iff $r = r(i)$.

As mentioned in the introduction, one motivation for considering such a heuristic is that if a C-SAP problem has a very big isolated optimum, our heuristic should be able to find it. This is stated in the following theorem.

Theorem 2.1 *Let $(\mathcal{R},\mathcal{T})$ be a C-SAP instance with $\bar{p}$ a feasible point in $\{0,1\}^{N\times K}$. For the extended instance $(\mathcal{R},\mathcal{T}\cup T')$ with $T' := \{(i,r(i))|i\in N, r(i)$ such that $\bar{p}_{ir(i)}=1\}$ and weight*

$$w_{T'} := c := \frac{2^{1/\alpha}\sum_{T\in\mathcal{T}} w_T}{\varepsilon^{(n-1)+|\mathcal{R}|\beta/\alpha+1/\alpha}},$$

any sequence $p^0, p^1 := \bar{F}(p^0) \equiv \Psi(F(p^0)),\dots$ with $p^0\in\Delta^\varepsilon$ converges to $\Psi(\bar{p})$.

Proof: Let $p := p^j$ and $p' := F(p)$ for some $j\geq 0$. Moreover, assume w.l.o.g. that $\bar{p}_{i1}=1$ for all $i\in N$. Then

$$p'_{ir} = u_{ir}\cdot p_{ir} \quad \text{with} \quad u_{ir} = \frac{\Gamma_{ir}(p)^\alpha\cdot\Theta_{ir}(p)^\beta}{\sum_{s=1}^k p_{is}\Gamma_{is}(p)^\alpha\cdot\Theta_{is}(p)^\beta}$$

We show that $u_{ir}\leq 1/2$ for all $i\in N, r\in K\setminus\{1\}$.
Therefore we need that for all $i\in N$, $\Gamma_{ir}(p)\leq\sum_{T\in\mathcal{T}} w_T$ for all $r\in K\setminus\{1\}$; $\Gamma_{i1}(p)\geq c\prod_{j\neq i}p_{j1}\geq c\varepsilon^{n-1}$ and $\Theta_{i1}(p)=\prod_{R:(i,1)\in R}(1-\prod_{(j,s)\in R\setminus(i,1)}p_{js})\geq\varepsilon^{|\mathcal{R}|}$. The last inequality is true, since $\bar{p}$ is feasible and therefore $(i,1)\in R$ implies that there exists $(j',s')\in R, s'\neq 1$. Hence $(1-\prod_{(j,s)\in R\setminus(i,1)}p_{js})\geq 1-p_{j's'}\geq\varepsilon$.
Note that for $i\in N, r\neq 1$

$$\begin{aligned} u_{ir} &= \frac{\Gamma_{ir}(p)^\alpha\cdot\Theta_{ir}(p)^\beta}{\sum_{s=1}^k p_{is}\Gamma_{is}(p)^\alpha\cdot\Theta_{is}(p)^\beta} \leq \frac{(\sum_{T\in\mathcal{T}} w_T)^\alpha}{\Gamma_{i1}(p)^\alpha\cdot\Theta_{i1}(p)^\beta p_{i1}} \\ &\leq \frac{(\sum_{T\in\mathcal{T}} w_T)^\alpha}{c^\alpha\varepsilon^{\alpha(n-1)}\varepsilon^{|\mathcal{R}|\beta}\varepsilon} \leq \frac{(\sum_{T\in\mathcal{T}} w_T)^\alpha}{2(\sum_{T\in\mathcal{T}} w_T)^\alpha} = \frac{1}{2}\end{aligned}$$

Since $u_{ir}\leq 1/2$ for all $i\in N, r\neq 1$ and $\Psi(p')_{ir}\leq p'_{ir}$ if $p'_{ir}>\varepsilon$ we have $p^{j+1}_{ir}\leq\frac{1}{2}p^j_{ir}$ or $p^{j+1}_{ir}=\varepsilon$, which ensures convergence of $p^0, p^1 := \bar{F}(p^0),\dots$ to $\Psi(\bar{p})$. ♣
Clearly the factor Γ in our dynamic is responsible for the nice behavior just discussed and we shall return later to the specific nature of the dynamic $F[\Gamma]$.
The above theorem shows that if the objective value of a feasible point $p\in\Delta^I$ is high enough, all trajectories lead to it. What happens, however, if an infeasible point p has such a high objective value? In fact, the repelling part of our operator acts nicely if p has a neighbor (with respect to change one value of one variable), for which less clauses in $\mathcal{R}$ are violated.

Theorem 2.2 *Let $\Gamma_{ir}\geq 1$ for $i\in N, r\in K$, $\beta>1$ and $\bar{p}\in\Delta^I$ be such that there exists a neighbor $q\in\Delta^I$ of $\bar{p}$ which satisfies strictly more clauses in $\mathcal{R}$ than $\bar{p}$. Then there exists a neighborhood U of $\Psi(\bar{p})$ such that $Round(\bar{F}(p))\neq Round(p), p\in U$.*

Proof: Note first that the condition $\Gamma_{ir}\geq 1$ is not restrictive, since for $\tilde{z} := z+\sum_{ir}p_{ir}$ we have $\tilde{z} := z+n, \forall p\in\Delta$ and $\tilde{z}$ satisfies the hypothesis on Γ. The proof is similar to that of Proposition 3.1 in [4] and we give here just the main ideas. Let $\bar{p}\in\Delta^I$ be an assignment for which u_1 clauses are not satisfied and q, obtained from $\bar{p}$ by changing the value of one decision variable, for which $u_2<u_1$ clauses are not satisfied. W.l.o.g. we assume that $\bar{p}_{i1}=1$, $\forall i$ and $q_{12}=1$. For simplicity we assume that all clauses have cardinality two. For $p=\Psi(\bar{p})$ those factors of $\Theta_{11}(p) := \prod_{R:(1,1)\in R}(1-\prod_{(j,s)\in R\setminus(1,1)}p_{js})$ corresponding to clauses satisfied by $\bar{p}$ are equal to $1-\varepsilon$ and their product can be approximated by 1 for our purpose (Actually we chose $\varepsilon>0$ small enough for our needs). The other factors (corresponding to violated clauses) are equal to ε and there are $u\leq u_1$ of them, so that

$\Theta_{11}(p) \approx \varepsilon^{u}$. Similarly, $\Theta_{12}(p) \approx \varepsilon^{u_2-(u_1-u)}$ and therefore $\Theta_{12}(p)/\Theta_{11}(p) \approx \varepsilon^{u_2-u_1}$. Let Γ_{max} be an upper bound for $\Gamma_{ir}, i \in N, r \in K$, then

$$\frac{F(p)_{12}}{F(p)_{11}} \geq \frac{\Theta_{12}^{\beta} p_{12}}{\Gamma_{max}^{\alpha} \Theta_{11}^{\beta} p_{11}} \geq \frac{\Theta_{12}^{\beta} \varepsilon}{\Gamma_{max}^{\alpha} \Theta_{11}^{\beta}} \approx \frac{1}{\Gamma_{max}^{\alpha}} \varepsilon^{\beta(u_2-u_1)+1}.$$

By a good choice of $\varepsilon > 0$ this can be made as large as needed to ensure that $F(p)_{12} \geq F(p)_{11}$. This can be done a priori for each instance of C-SAP (on the basis of the number of the variables, maximal size of the clauses, and upper bound for the ratio of the Γ_{ir} all assumed to be greater than one). By continuity, the result holds also in some $U(p)$. ♣

These two theorems describe nice properties of our combination of the "Γ-part" and the "Θ-part". If for example we construct a penalized objective function of the form

$$\tilde{z}(p) = z(p) - K(\sum_{R \in \mathcal{R}} \prod_{(i,r) \in R} p_{ir})$$

and maximize it by some gradient procedure, we loose these properties, if we want K to be large enough for a hill climbing procedure to prefer a feasible assignment with objective value 0 to any non feasible neighbor with value $c \gg 0$. In this respect a reasonable choice for K would be the sum of the coefficients of z. Consider now the $(n = 2,\ k = 2)$ instance $\max c p_{11} p_{21}$ s.t. $p_{11}p_{22} = 0,\ p_{12}p_{21} = 0$ and $p \in \Delta$. Our choice gives

$$\tilde{z}(p) = c p_{11} p_{21} - c(p_{11}p_{22} + p_{12}p_{21}).$$

The direction of the resulting gradient has then no dependency on c, as well as the basin of attraction of $(0,0)$ and $(1,1)$. Hence the property of Theorem 2.1 is lost.

By Theorem 2.2 we can not converge to an infeasible point having better neighbors, from the feasibility point of view. How about convergence in general? The following example shows that cycling is possible: For the problem

$$\begin{aligned} \max \quad & z(p) = p_{12} + p_{22} + 5000(p_{11}p_{21}) \\ \text{s.t.} \quad & p_{11}p_{21} = 0, \quad p \in \Delta^I \end{aligned}$$

consider F of (5) with $\alpha = 1$ and $\beta = 4$. Note that $\{p \in \Delta | p_{11} = p_{21}\}$ is invariant under F due to the symmetry of the instance and let $f(p_{11}) := F(p_{11}, 1 - p_{11}, p_{21} = p_{11}, 1 - p_{21})$. It can be shown that f has a cycle of period 3 ($f(0.174) \approx 0.9979$, $f(0.9979) \approx 4 \cdot 10^{-5}$, $f(4 \cdot 10^{-5}) \approx 0.174$) and therefore cycles of any period exists by Sarkovskii's theorem [5].

Nevertheless, in our numerical experiences convergence of the algorithm was observed and a useful stopping criterion can be given for testing if for each $i \in N$ the biggest component is going to increase for ever given that the fitness functions $\xi_{ir}, i \in N, r \in K$ are polynomials. We show it for functions ξ_{ir} with positive coefficients (the necessary modifications for negative coefficients is straightforward). The stopping criterion works as follows: Given $p \in \Delta^0$ (w.l.o.g. we assume that $p_{i1} > p_{ir},\ \forall r > 1$) we want to test if the sequence of iterates $F^t[\xi](p), t \geq 1$ converges to p^* where $p^*_{i1} = 1$ for all i. Decompose now for $i \in N, r \in K$ ξ_{ir} as $\xi_{ir} = \xi'_{ir} + \xi''_{ir}$ where ξ'_{ir} consists of those monomials in ξ_{ir} for which all variables have second index equal to 1. Note that $\xi_{ir}(p^*) = \xi'_{ir}(p^*)$. Finally let $q \in R^{N \times K}$

with $q_{i1} = 1$ and $q_{ir} = 1 - p_{ir}$, $\forall r > 1$.
Stopping criterion: If

$$\xi'_{i1}(p) > \xi'_{ir}(p^*) + \xi''_{ir}(q) \text{ for all } i \in N, r > 1 \text{ then STOP.} \tag{7}$$

Note simply that under the assumption that p_{i1} is not decreasing for $i \in N$, $\xi'_{i1}(p)$ is and will remain a lower bound for $\xi_{i1}(F[\xi]^t(p))$, $t \geq 1$ and for $i \in N, r > 1$, $\xi'_{ir}(p^*)+\xi''_{ir}(q)$ is and will remain an upper bound for $\xi_{ir}(F[\xi]^t(p))$ for $t > 1$. Therefore we have for all $t \geq 1$, $\xi_{i1}(F[\xi]^t(p)) > \xi'_{ir}(F[\xi]^t(p))$, $r > 1$ and $F[\xi]^t(p)_{i1}$ will remain greater than $F[\xi]^t(p)_{ir}$, $r > 1$. Moreover, if $(F[\xi]^t(p^0),\ t \geq 1)$ converges to $p^* \in \Delta^I$ and furthermore $\xi_{i1}(p^*) > \xi_{ir}(p^*)$ for all $i \in N$ and $r > 1$, simple continuity arguments show that there exist $n \geq 1$ and $p = F[\xi]^n(p^0)$ for which this test is successful.
We discuss now the stability of a fixed point p in the interior of Δ^ε for an operator $F[\xi]$ with the property that $\partial(\xi_{ir})/\partial(p_{ir}) \equiv 0$ for all $i \in N, r \in K$ (which we have in our approach): Since p is a fixed point, $\xi_{ir}(p) = N_i(p) = \sum_r p_{ir}\xi_{ir}(p)$. Moreover, $\partial(F[\xi]_{ir})/\partial(p_{ir}) = 1/N_i^2 \cdot (\xi_{ir}N_i - p_{ir}\xi_{ir}^2) = 1/N_i^2 \cdot (N_i^2(1-p_{ir})) = 1-p_{ir}$. It follows that the trace of the differential $DF(p)$ of $F[\xi]$ at p is $n(k-1)$, viewing $F[\xi]$ as a mapping with domain R^{nk}. Note that Δ^0 (actually the range of $F[\xi]$) has dimension $n(k-1)$ and that $DF(p)$ has by construction n vanishing eigenvalues. So the sum of the eigenvalues is at least the number of non vanishing eigenvalues. It follows that unless all non vanishing eigenvalues are equal to one, at least one of them will have modulus strictly greater than one, a criterion sufficient to imply that the fixed point p is unstable. The following slight modification of $F[\xi]$ implies this desirable property for all fixed points in the interior of Δ^ε [4]. Let $\Omega_{ir} := \prod_{1 \leq s \leq k, s \neq j} (1-p_{is})$ and take now as fitness function $\xi_{ir}\Omega_{ir}^h$ with $0 < h$. The trace of the corresponding differential is then $n(k-1)+nh$ and all fixed points in the interior are unstable. Furthermore, h can be chosen small enough to have otherwise a negligible effect and not alter our results in Δ^ε. This Ω-modification is a theoretical artefact and was not implemented.
To conclude this section, we present some results for the special case of C-SAP without feasibility constraints, i.e. $\mathcal{R} = \emptyset$.
Let V_z be the vector field associating to every $p \in \Delta^0$ the vector $V_z(p) := F[\Gamma](p) - p$.

Proposition 2.3 *Let Δ^0 be equipped with the Riemannian metric corresponding to the scalar product given by the $nk \times nk$ diagonal matrix $Q(p) = (q_{ir,js}(p))$ with $(q_{ir,ir}(p)) = \sum_r p_{ir}\Gamma_{ir}(p)/p_{ir}$, then V_z is the gradient of z with respect to Q.*

Proof: This metric is a straightforwad extension of the so-called Shahshahani metric (see for instance [8]). The proof consists in verifying directly that for every $p \in \Delta^0$ and for every vector v in the tangent space of Δ^0 at p (i.e $\sum_r v_{ir} = 0$ for all i) we have

$$V_z(p)'Q(p)v = Dz(p)(v)$$

where on the right-hand side we have the differential $Dz(p)$ of z at p applied to v (see [6] for the relevant background). ♣

An pleasant feature of our formulation (1) is that the objective is expressed as a harmonic function and therefore has no local optima in Δ^o if not constant. Recall on the other hand that gradient trajectories may not converge to a point. In our case however, it can be shown that if a trajectory has an accumulation point which is a local optima in the interior of a face, then it converges to it. Note furtermore that different functions on R^{nk} have the same restriction to Δ and induce different dynamics.

The introduction of our exponent α in (5) rests on the following generalization of [1] and can be proven similarly.

Proposition 2.4 *Let* $g : \mathbb{R}_+ \to \mathbb{R}_+$ *be monotone and concave and let* $G_{ir}(p) := g(\Gamma_{ir}(p))\ p \in \Delta^0,\ i \in N, r \in K$. *Then*

$$z(F[G](p)) \geq z(p)\ p \in \Delta^0$$

and $z(F[G](p)) = z(p)$ *only if* $F[G](p) = p$.

3. Implementation and empirical results

We tested our algorithm for two different problem classes. The first class is max-cut. The second contains randomly generated C-SAP instances for which finding feasible solutions is neither too easy nor too hard. In order to accelerate the evolution of the system, F of (5) is applied in a sequentially deterministic block Gauss-Seidel fashion, i.e. for $j \in N$ let

$$F_j(p)_{ir} := \begin{cases} \frac{p_{ir}\Gamma_{ir}(p)^\alpha \cdot \Theta_{ir}(p)^\beta}{\sum_{s=1}^k p_{is}\Gamma_{is}(p)^\alpha \cdot \Theta_{is}(p)^\beta} & i = j \\ p_{ir} & \text{otherwise} \end{cases} \tag{8}$$

Instead of F of (5) we apply $F' := F_n \circ \ldots \circ F_1$.

3.1 Empirical results for the max-cut problem

Given a graph $G = (V, E)$ and non-negative weights w_{ij} for $(i,j) \in E$. The max-cut problem consists in finding

$$\max_{S \subseteq V} \sum_{\substack{(i,j)\in E \\ i \in S, j \notin S}} w_{ij}$$

This problem can be formulated as a C-SAP with $n := |V|$ and $k = 2$ with: $p_{i1} = 1$ if node $i \in S$ and $p_{i2} = 1$ otherwise. The set of weighted clauses $\mathcal{T}$ is

$$\mathcal{T} := \cup_{(i,j)\in E} \left(\{(i_1, j_2)\} \cup \{(i_2, j_1)\}\right) \quad \text{and } \mathcal{R} = \emptyset.$$

The objective function of the C-SAP is therefore

$$z(p) = \sum_{(i,j)\in E} w_{ij}(p_{i1}p_{j2} + p_{j1}p_{i2}). \tag{9}$$

To solve the max-cut problem we use the operator F', where we choose $\alpha = 0.5$ and of course $\beta = 0$. Moreover, we use a slightly stronger stopping criterion as (7): for a strict local optima p^* a region, so-called guarenteed basin of attraction $GBA(p^*)$, can be determined, which is contained in the basin of attraction of p^*. Our algorithm runs as follows: we take the best solution found from 10 starting points with components randomly chosen from a given interval $J = [0.4995, 0.5005]$. For

each starting point we iterate $p^0, p^1 := F'(p^0), \ldots$ until a generated point, say p^k is either in the guaranteed basin of attraction $GBA(p^*_\nu)$ for some $p^*_\nu \in \Delta^I$ or for all $i \in N$ the biggest component $\max\{p^k_{i1}, p^k_{i2}\} \geq 1 - 10^{-5}$ or the number of iteration k attains a given iteration limit (= 1000).
We tested the algorithm on 10 randomly generated max cut instances with $n = 50$ vertices, edge probability of 0.2 and weights w_{ir} uniformly distributed in the interval $[1, 100]$ for $i \in N, r \in \{1, 2\}$. These instances have been solved optimally using an interior point algorithm in a branch and bound framework [3] and therefore the optima are known.
Each of the 10 instances has been solved 20 times (=20 runs). In all runs we were at most 2% off the optimum and on the average only 0.25%. Moreover, for all 10 problem instances the optimum has been found in the 20 runs (5 times at least for each instance, 11 times in the average). One run took about 5 seconds on a Pentium Pro 200 processor. Compared with the case where the GBA part of the stopping criteria is omitted the GBA criterion gives a speed up of about 80%.

3.2 Empirical results for a class of C-SAP

We tested our algorithm on randomly generated instances of C-SAP and compared it with the results of a Tabu Search (TS) [7]. This TS is in a working stage and possibly the results there can still be improved. The generated instances have $n = 100$ and $k = 5$ and $\mathcal{R}$ contains 10'000 randomly generated clauses of cardinality 3. Moreover $\mathcal{T}$ contains randomly generated clauses with cardinality between 1 and 5: 500 clauses of cardinality 1 with weights unifomly distributed in $[10, 20]$, 800 of cardinality 2 with weights in $[100, 200]$, 1000 of cardinality 3 with weights in $[500, 750]$, 800 of cardinality 4 with weights in $[750, 1500]$ and 200 of cardinality 5 with weights in $[1500, 2500]$. This choice was motivated by the following observations: first, there is no apparant relation between the best solutions found for this problem and the ones with $\mathcal{R} = \emptyset$; second, the mixture of the clauses in $\mathcal{T}$ satisfied by the best found solutions has no apparent pattern.

	DSH			Tabu		
	min	avg	max	min	avg	max
P1	42933	44914	46474	44436	46725	49809
P2	50329	52918	55491	51932	53449	55623
P3	52174	52795	54803	51564	54843	58549
P4	50117	51593	53886	53386	55526	59128
P5	51401	53959	56757	54968	57035	59849
P6	50687	52583	53960	50840	54159	57679
P7	49494	51631	54311	48233	53465	55897
P8	49353	52028	54162	50701	52428	54219
P9	50822	53273	56188	53226	54729	58295
P10	48670	50454	51848	51371	54895	57575

Table 1: Comparison DSH with Tabu

For the implemented algorithm (DSH), we take the best solution found from 100 starting points (this number has been chosen in order to use equal CPU-time to TS from one starting point.) For

each starting point we essentially iterate $p^0, p^1 := \Psi(F(p)), \ldots, p_{200}$. In order not to get stuck at the boundary of Δ^ε we recenter the points every 10 iterations (i.e. add a constant to all variables and normalize). Moreover, the encountered solutions will be improved by a greedy algorithm (first priority: satisfyability, second priority: objective value).
Table 1 shows the results of our heuristic DSH compared with TS for 10 instances. Each instance was solved 10 times (= 10 runs) and the minimum (min), average (avg) and maximum (max) over these 10 runs are given in the table for the 10 instances. Our heuristic gives slightly worse results compared to Tabu (for the average 3.7%). However, we are confident that there is still some room for improvement, in particular the order of the variables in our Gauss-Seidel procedure could take advantage of the available information. Moreover, DSH is less sensitive to shortage of the available CPU time. To conclude, DSH is not the best heuristic, but it is an interesting alternative which could be combined with other LS heuristics, due to its completely different approach.

Thanks: This work was partly supported by Fond National, Switzerland.

Corresponding Author: M. Cochand,
Institute for Operations Research,
Swiss Federal Institute of Technology,
ETH-Zentrum, 8092 Zürich, SWITZERLAND

References

[1] Leonard E. Baum and George R. Sell. Growth transformations for functions on manifolds. *Pacific Journal of Mathematics*, 27(2), 1968.

[2] Immanuel M. Bomze. Evolution towards the Maximum Clique. *Journal of Global Optimization*, 10,pp 143-164, 1997.

[3] Michael Burkard An interior point algorithm for solving Max Cut problems *Diploma thesis, Technical University of Graz*, April,1994.

[4] Maurice Cochand. A fixed point operator for the generalized maximum satisfiability problem. *Discrete Applied Mathematics*, 46:117–132, 1993.

[5] Robert L. Devaney. *An introduction to Chaotic Dynamical Systems.* The Benjamin/Cummings Publishing Company, 1986.

[6] Uwe Helmke and John B. Moore. *Optimization and Dynamical Systems.* Springer, 1994.

[7] Alain Hertz, Daniel Kobler. A Tabu Search Algorithm for the Constrained Semi-Assignment Problem *in preparation.*

[8] Josef Hofbauer and Karl Sigmund. *The theory of Evolution and Dynamical Systems.* Cambridge University Press, 1988.

[9] Jürgen W. Weibull. *Evolutionary Game Theory.* The MIT Press, 1995.

[10] Wing Sing Wong. Gradient Flows for Local Minima of Combinatorial Optimization Problems. *Fields Institute Communications*, 3, 1994.