The Deviation Constraint

Pierre Schaus[1], Yves Deville[1], Pierre Dupont[1], Jean-Charles Régin[2]

[1] Université catholique de Louvain, Belgium
{pschaus, yde, pdupont}@info.ucl.ac.be
[2] ILOG, France
regin@ilog.fr

Abstract. This paper introduces `DEVIATION`, a soft global constraint to obtain balanced solutions. A violation measure of the perfect balance can be defined as the L_p norm of the vector variables minus their mean. `SPREAD` constraints the sum of square deviations to the mean [5, 7] *i.e.* the L_2 norm. The L_1 norm is considered here. Neither criterion subsumes the other but the design of a propagator for L_1 is simpler. We also show that a propagator for `DEVIATION` runs in $\mathcal{O}(n)$ (with respect to the number of variables) against $\mathcal{O}(n^2)$ for `SPREAD`.

1 Introduction

We consider the Balanced Academic Curriculum Problem (BACP) [1] as a motivating example. The goal is to assign a period to each course in a way that the prerequisite relationships are satisfied and *the academic load of each period is balanced*. This last constraint makes BACP a Constraint Optimization Problem where the objective is to maximize the balancing property.

In BACP the mean m of a solution is a constant of the problem since the load of each course and the number of periods are given. A hard balancing constraint would impose all periods to take a same load m. This often results in an over-constrained problem without solution. For a set of variables $\mathcal{X} = \{X_1, X_2, ..., X_n\}$ and a given fixed mean m, a violation measure of the perfect balance property can be defined as the L_p-norm of the vector $[\mathbf{X} - \mathbf{m}]$ with $\mathbf{X} = [X_1, X_2, ..., X_n]$, $\mathbf{m} = [m, m, ..., m]$ such that $\sum_{i=1}^{n} X_i = n.m$. The L_p-norm of $[\mathbf{X} - \mathbf{m}]$ is defined as $(\sum_{i=1}^{n} |X_i - m|^p)^{\frac{1}{p}}$ with $p \geq 0$.

Following the scheme proposed by Régin et al. [6] to soften global constraints, we define a violation of the perfect balance constraint as a cost variable L_p in the global balance constraint: `soft-balance`$(\mathcal{X}, m, L_p)$ constraint holds if and only if L_p-norm$([\mathbf{X} - \mathbf{m}]) = L_p$ and $\sum_{i=1}^{n} X_i = n.m$.

The interpretation of the violation to the mean for some specific norms is given below.

- L_0: $|\{X \in \mathcal{X} | X \neq m\}|$ is the number of values different from the mean.
- L_1: $\sum_{X \in \mathcal{X}} |X - m|$ is the sum of deviations from the mean.
- L_2: $\sum_{X \in \mathcal{X}} (X - m)^2$ is the sum of square deviations from the mean.
- L_∞: $\max_{X \in \mathcal{X}} |X - m|$ is the maximum deviation from the mean.

Note that none of these balance criteria subsumes the others. For instance, the minimization of L_1 does not imply in general a minimization of criterion L_2. This is illustrated on the following example. Assume a constraint problem with four solutions given in Table 1. The most balanced solution depends on the chosen norm. Each solution exhibits a mean of 100 but each one optimizes a different norm.

sol. num.	solution	L_0	L_1	L_2	L_∞
1	100 100 100 100 30 170	**2**	140	9800	70
2	60 80 100 100 120 140	4	**120**	4000	40
3	70 70 90 110 130 130	6	140	**3800**	30
4	71 71 71 129 129 129	6	174	5046	**29**

Table 1. Illustration showing that no balance criterion defined by the norm L_0, L_1, L_2 or L_∞ subsumes the others. The smallest norm is indicated in bold character. For example, solution 2 is the most balanced according to L_1.

The norm L_∞ has already been used in two previous works [2, 4] to solve BACP. `SPREAD` is a constraint for L_2 [5, 7]. A constraint for L_0 can easily be implemented using an $\texttt{ATLEAST}(i, [X_1, ..., X_n], m)$ constraint for $|\{X \in \mathcal{X} | X \neq m\}| \leq i$ and a $\texttt{SUM}([X_1, ..., X_n], n.m)$ constraint to ensure a mean of m. In this paper, a global constraint and its propagators for the L_1 norm with fixed mean is presented. This constraint is formulated in the following definition:

Definition 1. *A set of finite domain integer variables $\mathcal{X} = \{X_1, X_2, ..., X_n\}$, one mean value m and one interval variable D are given.*

The constraint $\texttt{DEVIATION}(\mathcal{X}, m, D)$ states that the collection of values taken by the variables of $\mathcal{X}$ exhibits an arithmetic mean m and a sum of deviations to m of D. More formally, $\texttt{DEVIATION}(\mathcal{X}, m, D)$ holds if and only if

$$n.m = \sum_{i=1}^{n} X_i \quad \text{and} \quad D = \sum_{i=1}^{n} |X_i - m|.$$

For the constraint to be consistent, $n \times m$ must be an integer. As a consequence $n \times D$ is also an integer.

Outline of the paper:

Section 2 mainly reviews preliminaries notions relative to constraint programming such as filtering, domain-consistency and bound-consistency. We also define some useful notations. Section 3 motivates the need of a global filtering algorithm for `DEVIATION` in terms of filtering. Section 4 explains the propagators narrowing the domain of D. This filtering makes use of the minimization and maximization of the sum of deviations. The minimization is solved in linear time. The maximization is proved to be $\mathcal{NP}$-complete; however, an approximated upper bound can be calculated in linear time as well. Section 5 describes the filtering algorithm from m and D to the variables $\mathcal{X}$. The idea is similar to a bound consistency filtering algorithm for a `SUM` constraint but including the sum of deviations constraint. Section 6 shows that our propagators do not achieve bound-consistency. Section 7 gives a relaxation of `SPREAD` with `DEVIATION`. Finally, Section 8 evaluates the efficiency of the presented propagators in terms of filtering on randomly generated instances.

2 Background and notations

Basic constraint programming concepts largely inspired from Section 2 of [8] are introduced.

Let X be a finite-domain (discrete) *variable*. The *domain* of X is a set of ordered values that can be assigned to X and is denoted by $Dom(X)$. The minimum (resp. maximum) value of the domain is denoted by $X^{\min} = \min(Dom(X))$ (resp. $X^{\max} = \max(Dom(X))$. Let $\mathcal{X} = \{X_1, X_2, ..., X_k\}$ be a sequence of variables. A *constraint* C on $\mathcal{X}$ is defined as a subset of the Cartesian product of the domains of the variables in $\mathcal{X}$: $C \subseteq Dom(X_1) \times Dom(X_2) \times ... \times Dom(X_k)$. A tuple $(v_1, ..., v_k) \in C$ is called a *solution* to C. A value $v \in Dom(X_i)$ for some $i = 1, ..., k$ is *inconsistent* with respect to C if it does not belong to a tuple of C, otherwise it is *consistent*. C is inconsistent if it does not contain a solution. Otherwise, C is called consistent.

A constraint satisfaction problem, or a CSP, is defined by a finite sequence of variables $\mathcal{X} = \{X_1, X_2, ..., X_n\}$, together with a finite set of constraint $\mathcal{C}$ each on a subset of $\mathcal{X}$. The goal is to find an *assignment* $X_i := v$ with $v \in Dom(X_i)$ for $i = 1, ..., n$, such that all constraints are satisfied. This assignment is called a solution to the CSP.

The solution process of constraint programming interleaves constraint *propagation*, or propagation in short, and search. The search process essentially consists of enumeration all possible variable-value combinations, until a solution is found or it is proved that none exists. We say that this process constructs a search tree. To reduce the exponential number of combinations, constraint propagation is applied to each node of the search tree: Given the current domains and a constraint C, the propagator for C removes domain values that do not belong to a solution to C. This is repeated for all constraints until no more domain value can be removed. The removal of inconsistent domain values is called *filtering*.

In order to be effective, filtering algorithms should be efficient, because they are applied many times during the search process. They should furthermore remove as many inconsistent values as possible. If a filtering algorithm for a constraint C removes all inconsistent values from the domains with respect to C, we say that it makes C *domain-consistent*. It is possible to achieve domain-consistency in polynomial time for some constraints such as `AllDiff` but for other constraints such as `SUM` this would be too costly. In such cases a weaker notion of consistency called *bounds-consistency* (also called interval-consistency) appears to be highly cost-effective. A constraint C is bound consistent if the bounds of the domain of each variable implied in C belongs to at least one solution of C. The idea it to bound the domain of each variable by an interval and make sure that the end-points of the intervals obey the domain-consistency requirement. If not, the upper and lower bounds of the intervals can be tightened until bounds-consistency is achieved.

Proposition 1 states that achieving domain-consistency for `DEVIATION` is not more difficult than for arithmetic constraints in general.

Proposition 1. *Achieving domain-consistency for* `DEVIATION` *is* $\mathcal{NP}$*-Complete.*

Proof. In the particular case where $Dom(D) = [0, +\infty]$, achieving domain-consistency for `SUM` constraint reduces to achieving domain-consistency for `DEVIATION`. This is $\mathcal{NP}$-Complete since the *subset sum* problem (see [3]) can be reduced to achieving domain-consistency for the `SUM` constraint. ■

Domains of variables are considered as full and can be described by the interval $Dom(X) = [X^{\min}..X^{\max}]$.

Definition 2 introduces some useful notations and can be easily understood with the graphical illustration on Figure 1. A numerical example is also given in Example 1.

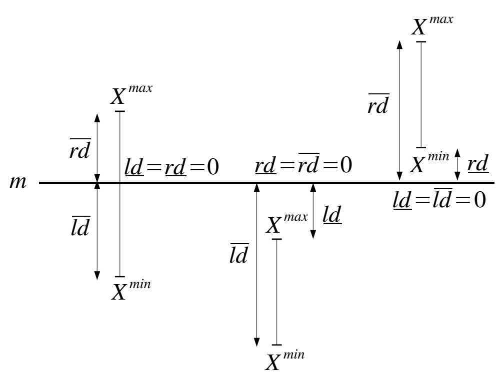

Fig. 1. Illustration of the definitions introduced in Definition 2 for three variables.

Definition 2. *For a variable X and a given value m, the upper bounds on the right and left deviation are respectively*

- $\overline{rd}(X, m) = \max(0, X^{\max} - m)$ *and*
- $\overline{ld}(X, m) = \max(0, m - X^{\min})$.

The sum of these values over $\mathcal{X}$ are respectively

- $\overline{RD}(\mathcal{X}, m) = \sum_{X \in \mathcal{X}} \overline{rd}(X, m)$ *and*
- $\overline{LD}(\mathcal{X}, m) = \sum_{X \in \mathcal{X}} \overline{ld}(X, m)$.

The same idea holds for the lower bounds on the deviations:

- $\underline{rd}(X, m) = \max(0, X^{\min} - m)$.
- $\underline{ld}(X, m) = \max(0, m - X^{\max})$.
- $\underline{RD}(\mathcal{X}, m) = \sum_{X \in \mathcal{X}} \underline{rd}(X, m)$.
- $\underline{LD}(\mathcal{X}, m) = \sum_{X \in \mathcal{X}} \underline{ld}(X, m)$.

For a variable $X_i \in \mathcal{X}$ *we define:*

- $\overline{LD}_i(\mathcal{X}, m) = \overline{LD}(\mathcal{X}, m) - \overline{ld}(X_i, m)$ *and*
- $\overline{RD}_i(\mathcal{X}, m) = \overline{RD}(\mathcal{X}, m) - \overline{rd}(X_i, m)$.

To alleviate notations, $(\mathcal{X}, m)$ *are sometimes omitted. For example* $\overline{LD}(\mathcal{X}, m)$ *is simply written* $\overline{LD}$.

Example 1. Let $\mathcal{X} = \{X_1, X_2, X_3, X_4\}$ be four variables with domains $Dom(X_1) = [8, 10]$, $Dom(X_2) = [4, 7]$, $Dom(X_3) = [1, 5]$ and $Dom(X_4) = [3, 4]$. The following table exhibits the quantities introduced in Definition 2.

i	$\overline{rd}(X_i, 5)$	$\overline{ld}(X_i, 5)$	$\underline{rd}(X_i, 5)$	$\underline{ld}(X_i, 5)$
1	5	0	3	0
2	2	1	0	0
3	0	4	0	0
4	0	2	0	1
$\sum_i$	**7**	**7**	**3**	**1**
	$\overline{RD}_i(\mathcal{X}, 5)$	$\overline{LD}_i(\mathcal{X}, 5)$	$\underline{RD}_i(\mathcal{X}, 5)$	$\underline{LD}_i(\mathcal{X}, 5)$
1	2	7	0	1
2	5	6	3	1
3	7	3	3	1
4	7	5	3	0

The filtering for `DEVIATION` is based on the next theorem stating that the sum of deviations above and under the mean are equal.

Theorem 1. *Let* $\mathcal{X} = \{X_1, ..., X_n\}$*. The equality* $n.m = \sum_{X \in \mathcal{X}} X$ *holds if and only if* $\sum_{X > m}(X - m) = \sum_{X < m}(m - X)$.

Proof. $n.m = \sum_{X \in \mathcal{X}} X$ can be rewritten $0 = \sum_{X \in \mathcal{X}} X - n.m = \sum_{X > m}(X - m) + \sum_{X < m}(X - m) + \sum_{X = m}(X - m) = \sum_{X > m}(X - m) - \sum_{X < m}(m - X)$. ∎

Property 1 *Let* $\mathcal{X} = \{X_1, ..., X_n\}$*. An assignment on* $\mathcal{X}$ *satisfies:*

- $\sum_{X > m}(X - m) \in [\underline{RD}(\mathcal{X}, m), \overline{RD}(\mathcal{X}, m)]$ *and*
- $\sum_{X < m}(X - m) \in [\underline{LD}(\mathcal{X}, m), \overline{LD}(\mathcal{X}, m)]$.

Theorem 2. *`DEVIATION`*$(\mathcal{X}, m, D)$ *is consistent only if the following conditions are satisfied:*

1. $\underline{RD}(\mathcal{X}, m) \leq \frac{D^{\max}}{2}$
2. $\underline{LD}(\mathcal{X}, m) \leq \frac{D^{\max}}{2}$
3. $\overline{RD}(\mathcal{X}, m) \geq \frac{D^{\min}}{2}$
4. $\overline{LD}(\mathcal{X}, m) \geq \frac{D^{\min}}{2}$
5. $\left[\underline{LD}(\mathcal{X}, m), \overline{LD}(\mathcal{X}, m)\right] \cap \left[\underline{RD}(\mathcal{X}, m), \overline{RD}(\mathcal{X}, m)\right] \neq \phi$

Proof. 1. If $\underline{RD}(\mathcal{X}, m) > \frac{D^{\max}}{2}$ then $\sum_{X>m}(X - m) > \frac{D^{\max}}{2}$ (Property 1). Hence $\sum_{i=1}^{n} |X_i - m| > D^{\max}$ (by Theorem 1).
2., 3. and 4. similar to 1.
5. Direct consequence of Theorem 1 and Property 1. ■

3 Naive implementation

This section explains why a naive implementation of `DEVIATION` by decomposition into more elementary constraints is not optimal in terms of filtering.

As stated in Definition 1, `DEVIATION`$(\mathcal{X}, m, D)$ holds if and only if $n.m = \sum_{i=1}^{n} X_i$ and $D = \sum_{i=1}^{n} |X_i - m|$. This suggests a natural implementation of the constraint by decomposing it into two `SUM` constraints. Figure 2 illustrates that the filtering obtained with the decomposition is not optimal.

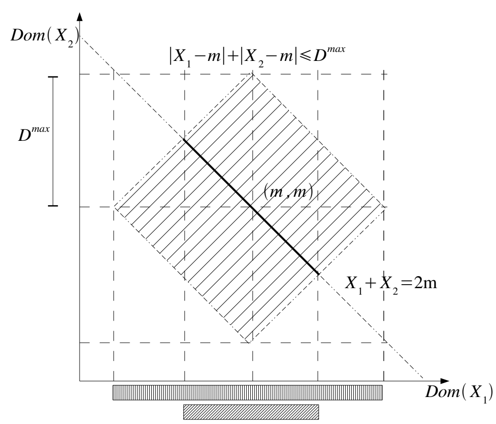

Fig. 2. Filtering of X_1 with decomposition and with `DEVIATION`.

Assume two variables X_1, X_2 with unbounded finite domains and the constraint

$$\texttt{DEVIATION}(\{X_1, X_2\}, m, D \in [0, D^{\max}]).$$

The diagonally shaded square (see Figure 2) delimits the set of points such that $|X_1 - m| + |X_2 - m| \leq D^{\max}$. The diagonal line is the set of points such that $X_1 + X_2 = 2.m$. The unbounded domains for X_1 and X_2 are bound-consistent for the mean constraint. The vertically shaded rectangle defines the domain of

X_1 after a bound-consistent filtering for $|X_1 - m| + |X_2 - m| \leq D^{\max}$. The set of solutions for DEVIATION is the bold diagonal segment obtained by intersecting the square surface and the diagonal line. It can be seen on the figure that more filtering is possible. Bound consistent filtering on the bold diagonal segment leads to a domain of X_1 defined by the diagonally shaded rectangle. In conclusion a bound consistent filtering for the decomposition leads to $Dom(X_1) = Dom(X_2) = [m - D^{\max}, m + D^{\max}]$ while a bound consistent filtering for DEVIATION$(\{X_1, X_2\}, m, D \in [0, D^{\max}])$ leads to domains two times smaller $Dom(X_1) = Dom(X_2) = [m - \frac{D^{\max}}{2}, m + \frac{D^{\max}}{2}]$.

4 Filtering of D

The filtering of D to achieve bound consistency requires to solve two optimization problems: the minimization and maximization of the sum of deviations from a given mean m.

Definition 3. *$\underline{D}$ and $\overline{D}$ denote the optimal values to problems:*

$$\underline{D} = \min \sum_{i=1}^{n} |X_i - m| \quad \text{and} \quad \overline{D} = \max \sum_{i=1}^{n} |X_i - m|$$

$$\text{such that :} \sum_{i=1}^{n} X_i = n.m$$

$$X_i^{\min} \leq X_i \leq X_i^{\max},\ 1 \leq i \leq n$$

$$X_i \in \mathbb{Q},\ 1 \leq i \leq n.$$

Values $\underline{D}$ and $\overline{D}$ can be used to filter the domain of D:

$$Dom(D) \longleftarrow Dom(D) \cap [\underline{D}, \overline{D}].$$

The remaining of this section mainly explains how $\underline{D}$ can be computed in linear time with respect to the number of variables $n = |\mathcal{X}|$. Unfortunately finding $\overline{D}$ is an $\mathcal{NP}$-complete problem and the best we can do is to design a good upper bound for it as explained at the end of this section.

Definition 4 characterizes an optimal solution to the problem of finding $\underline{D}$.

Definition 4 (up and down centered assignment). *Let $\mathcal{X} = \{X_1, ..., X_n\}$. Let $A : X \rightarrow Dom(X)$ be an assignment of $X \in \mathcal{X}$. The quantity $s(A)$ denotes the sum of assigned values: $s(A) = \sum_{X \in \mathcal{X}} A(X)$.*

*An assignment A is said to be **up-centered** when:*

$$A(X) \begin{cases} = X^{\min} & \text{if } X^{\min} \geq s(A)/n \\ \leq s(A)/n & \text{otherwise} \end{cases}$$

In other words, each variable with minimum domain value larger than the mean of the assigned values takes its minimum domain value and the other variables take values smaller than the mean of the assigned values.

An assignment A is said to be ***down-centered*** *when:*

$$A(X)\begin{cases} = X^{\max} & \text{if } X^{\max} \le s(A)/n \\ \ge s(A)/n & \text{otherwise} \end{cases}$$

In other words, each variable with maximum domain value smaller than the mean of the assigned values takes its maximum domain value and the other variables take values larger than the mean of the assigned values.

Example 2. Considering the variables and domains of Example 1, the following assignment is up-centered with mean $17/4$*:*

$$A(X_1) = 8, A(X_2) = 4, A(X_3) = 2, A(X_4) = 3.$$

Theorem 3. *An assignment is an optimal solution to the problem of finding* $\underline{D}$ *if and only if it is a down-centered assignment or an up-centered assignment of mean* m*.*

Proof. (if) Given an assignment A of mean m *i.e.* $s(A) = n.m$. The only way to decrease the sum of deviations while conserving the mean m is to find a pair of variables X_i, X_j such that $A(X_i) > m, A(X_i) > X_i^{\min}, A(X_j) < m, A(X_j) < X_j^{\max}$ and to decrease $A(X_i)$ and increase $A(X_j)$ by the same quantity to make them closer to m. By definition of a left and down centered assignment, it is impossible to find such a pair X_i, X_j. Hence, up-centered and a down-centered assignments are optimal solutions.

(only if) Assume an assignment A neither down-centered nor up-centered such that $s(A) = n.m$. It is possible to find at least two variables $X_i, X_j \in \mathcal{X}$. One with $A(X_i) > m$ and $A(X_i) > X_i^{\min}$ (violation of up-centered) and one with $A(X_j) < m$ and $A(X_j) < X_j^{\max}$ (violation of down-centered). Let define $\delta = \min\left(A(X_i) - \max(X_i^{\min}, m), \min(X_j^{\max}, m) - A(X_j)\right)$. The assignment $A(X)$ is not optimal since the sum of deviations can be decreased by 2δ by modifying the assignment on X_i and X_j: $A'(X_i) = A(X_i) - \delta$ and $A'(X_j) = A(X_j) + \delta$. ■

Theorem 4. *If* `DEVIATION` *is consistent then*

$$\underline{D} = 2.\max(\underline{LD}(\mathcal{X}, m), \underline{RD}(\mathcal{X}, m)).$$

Proof. Assume $\underline{LD} \ge \underline{RD}$, then it is possible to build a down-centered assignment A of mean m and which is optimal by Theorem 3. For this assignment $\sum_{A(X)<m}(m - A(X)) = \underline{LD}$ (by Definition 2 of $\underline{LD}$). Since $\sum_{X>m}(X - m) = \sum_{X<m}(m - X)$ (by Theorem 1), the sum of deviations for this down-centered assignment is $\sum_{X\in\mathcal{X}} |A(X) - m| = 2.\underline{LD}$.

The case $\underline{LD} \le \underline{RD}$ is similar. The assignment is up-centered instead of down-centered. ■

Example 3. The variables and domains considered here are the same as those in Example 1. A mean $m = 5$ is considered. Using the computed values $\underline{LD}(\mathcal{X}, 5) = 1$ and $\underline{RD}(\mathcal{X}, 5) = 3$ from Example 1, it can be deduced that $\underline{D} = 2.\max(1, 3) = 6$.

Consequently, filtering on D for $\texttt{DEVIATION}(\mathcal{X} = \{X_1, X_2, X_3, X_4\}, m = 5, D \in [0, 7])$ leads to $Dom(D) = [6, 7]$.

Theorem 5. *Computing $\overline{D}$ is $\mathcal{NP}$-complete.*

Proof. It is possible to reduce the subset sum problem [3] to the problem of computing $\overline{D}$ (see Definition 3). This problem is not more difficult than the particular case where $m = 0$:

$$\overline{D} = \max \sum_{i=1}^{n} |X_i|$$
$$\text{such that :} \sum_{i=1}^{n} X_i = 0$$
$$X_i^{\min} \leq X_i \leq X_i^{\max},\ 1 \leq i \leq n.$$

Given a set of n positive values $\{b_1, ..., b_{n-1}, T\}$, the subset sum problem consists in finding if there exists a set of binary values $\{y_1, ..., y_{n-1}\}, y_i \in \{0,1\}, 1 \leq i < n$ such that $\sum_{i=1}^{n-1} y_i.b_i = T$. The reduction is the following:

- $X_i^{\min} = -\frac{b_i}{2}$ and $X_i^{\max} = \frac{b_i}{2}$ for $1 \leq i < n$.
- $X_n = \frac{\sum_{i=1}^{n-1} b_i}{2} - T$.
- There is a solution to the subset sum problem if $\overline{D} \geq \frac{\sum_{i=1}^{n-1} b_i}{2} + \left|\frac{\sum_{i=1}^{n-1} b_i}{2} - T\right|$. This constraint on the optimal value ensures that the optimal solution is such that $X_i \in \{-\frac{b_i}{2}, \frac{b_i}{2}\}$. The solution to the subset sum problem is then given by $y_i = 1$ if $X_i = \frac{b_i}{2}$ and $y_i = 0$ if $X_i = \frac{-b_i}{2}$. ■

Unless $\mathcal{P} = \mathcal{NP}$, the problem of computing $\overline{D}$ is exponential (Theorem 5). As explained in Section 2, in order to be effective, filtering algorithms should be as efficient as possible because they are applied many times during the search process. This is why we prefer to find efficiently a good upper bound $\overline{D}^{\uparrow}$ for $\overline{D}$ than to find its exact value. An upper bound which can be computed in $\mathcal{O}(n)$ is

$$\overline{D}^{\uparrow} = \sum_{i=1}^{n} \max\left(|X_i^{\max} - m|, |X_i^{\min} - m|\right).$$

The filtering on $Dom(D)$ becomes:

$$Dom(D) \longleftarrow Dom(D) \cap [\underline{D}, \overline{D}^{\uparrow}]$$

5 Filtering on $\mathcal{X}$

Two propagators could be considered to filter the domain of $\mathcal{X}$:

1. from $D^{\min}$ and m to $\mathcal{X}$ and
2. from $D^{\max}$ and m to $\mathcal{X}$.

Achieving bound-consistency for the first propagator is $\mathcal{NP}$-complete. Indeed, checking the consistency of one value requires to maximize the sum of deviations which is an $\mathcal{NP}$-complete problem (Theorem 5). The decomposition of `DEVIATION` presented in Section 3 can however be used to realize bound-consistency on constraint $\sum_{X \in \mathcal{X}} |X - m| \geq D^{\min}$. In any case, this filtering is useless if one seeks a balanced solution on $\mathcal{X}$. Hence, the remaining of this section focuses on a linear time filtering algorithm for the second propagator.

The filtering is based on the computation of the values $\overline{X}_i$ and $\underline{X}_i$ introduced in Definition 5.

Definition 5. *$\overline{X}_i$ and $\underline{X}_i$ are the optimal values to the following problems:*

$$\overline{X}_i = \max(X_i) \quad \text{and} \quad \underline{X}_i = \min(X_i)$$

$$\text{such that}: \sum_{j=1}^{n} X_j = n.m \tag{1}$$

$$\sum_{j=1}^{n} |X_j - m| \leq D^{\max} \tag{2}$$

$$X_j^{\min} \leq X_j \leq X_j^{\max}, 1 \leq j \leq n, j \neq i$$

$$X_j \in \mathbb{Q}, 1 \leq j \leq n.$$

The filtering rule on the domain of X_i can be simply written:

$$Dom(X_i) \longleftarrow Dom(X_i) \cap [\underline{X}_i, \overline{X}_i] \tag{3}$$

Theorem 6. *For a variable X_i, assuming the constraint is consistent, the following equalities hold:*

$$\overline{X}_i = \min\left(\frac{D^{\max}}{2}, \overline{LD}_i(\mathcal{X}, m)\right) - \underline{RD}_i(\mathcal{X}, m) + m.$$

$$\underline{X}_i = -\min\left(\frac{D^{\max}}{2}, \overline{RD}_i(\mathcal{X}, m)\right) + \underline{LD}_i(\mathcal{X}, m) + m.$$

Proof. Only $\overline{X}_i$ is considered because the proof for $\underline{X}_i$ is symmetrical with respect to m. Two cases can be considered:

- $\overline{LD}_i \leq \frac{D^{\max}}{2}$: By Theorem 1 the deviation above the mean and under the mean must be equal. Hence the optimal solution is such that $\overline{X}_i - m + \underline{RD}_i = \overline{LD}_i$. Constraint (2) is not tight in this case.
- $\overline{LD}_i > \frac{D^{\max}}{2}$: By Theorem 1 the constraint (1) means that the deviation above the mean and under the mean must be equal. The conjunction of constraint (1) with constraint (2) means that the deviation above and under the mean are equal and at most $D^{\max}/2$. Hence the optimal solution is such that $\overline{X}_i - m + \underline{RD}_i = \frac{D^{\max}}{2}$. Constraint (2) is tight in this case.

If both cases are considered together, equality $\overline{X}_i - m + \underline{RD}_i = \min(\frac{D^{\max}}{2}, \overline{LD}_i)$ holds at the optimal solution. ■

The filtering procedure on $\mathcal{X}$ applies rule (3) once on each $X_i \in \mathcal{X}$. This can be achieved in linear time with respect to the number of variables.

Example 4. Variables and domains considered are the same as in Example 1. The constraint considered is $\texttt{DEVIATION}(\mathcal{X} = \{X_1, X_2, X_3, X_4\}, m = 5, D \in [0, 7])$. Values $\overline{X}_i$ and $\underline{X}_i$ are: $\overline{X}_1 = \min(3.5, 7) - 0 + 5 = 8.5$, $\overline{X}_2 = \min(3.5, 6) - 3 + 5 = 5.5$, $\overline{X}_3 = \min(3.5, 3) - 3 + 5 = 5$, $\overline{X}_4 = \min(3.5, 5) - 3 + 5 = 5.5$, $\underline{X}_1 = -\min(3.5, 2) + 1 + 5 = 4$, $\underline{X}_2 = -\min(3.5, 5) + 1 + 5 = 2.5$, $\underline{X}_3 = -\min(3.5, 7) + 1 + 5 = 2.5$ and $\underline{X}_4 = -\min(3.5, 7) + 0 + 5 = 1.5$. Hence filtering rule (3) leads to filtered domains: $Dom(X_1) = [8, 8]$, $Dom(X_2) = [4, 5]$, $Dom(X_3) = [3, 5]$ and $Dom(X_4) = [3, 4]$.

6 Bound consistency for DEVIATION

The total filtering is achieved in $\mathcal{O}(n)$ as follows.

1. Filtering from $D^{\max}$ and m to $\mathcal{X}$: $\forall X \in \mathcal{X}, Dom(X) \longleftarrow Dom(X) \cap [\underline{X}, \overline{X}]$.
2. Filtering from $\mathcal{X}$ and m to D: $Dom(D) \longleftarrow Dom(D) \cap [\underline{D}, \overline{D}^{\uparrow}]$.

Even if it was possible to compute $\overline{D}$ efficiently, bound consistency is not necessarily obtained. The reason is that values $\overline{X}$, $\underline{X}$ (Definition 5) and $\underline{D}$ (Definition 3) are computed making the assumption that interval domains are defined on rational numbers $\mathbb{Q}$ rather than on integers $\mathbb{Z}$. As the next example shows, this can lead to miss some possible filtering.

Example 5. Assume a set of 10 variables $\mathcal{X}$ with domain $[0, 1]$ and a mean of $m = 0.5$. Theorem 4 gives a value $\underline{D} = 0$ because every domain overlaps m. In fact the only way to obtain an assignment respecting the mean constraint is to have five variables assigned to 0 and five to 1. For such an assignment, the minimal sum of deviations from the mean is 5 and not 0. Consequently, the constraint $\texttt{DEVIATION}(\mathcal{X}, m = 0.5, D \in [0, 3])$ is inconsistent but such an inconsistency will not be detected by our propagator.

Example 5 shows that all inconsistencies are not detected by the propagator. This occurs when the mean is not an integer but a rational number and when the domains of some variables include the mean. When the mean is an integer, such a problem does not occur and the propagator is bound-consistent.

7 Relation between SPREAD and DEVIATION

This section shows that DEVIATION can be used as a relaxation of SPREAD. This relaxation might be useful since the propagator of the former runs in $\mathcal{O}(n)$ against $\mathcal{O}(n^2)$ for the latter. Furthermore, DEVIATION is easier to implement.

The relaxation is illustrated graphically and the parameters given to DEVIATION to obtain the strongest relaxation as possible are expressed as a function of SPREAD parameters.

SPREAD($\mathcal{X}, m, \Delta^2$) holds if the m value is the average over $\mathcal{X}$ and the sum of square deviation to m is Δ^2. More formally SPREAD holds if $\sum_{x \in \mathcal{X}} X = n.m$ and $\sum_{x \in \mathcal{X}} (X - m)^2 = \Delta^2$.

From a geometrical point of view, $\sum_{x \in \mathcal{X}} (X - m)^2 \leq (\Delta^{\max})^2$ defines an hyper-sphere centered on $[m, ..., m]$ of radius $\Delta^{\max}$. The set of points satisfying $\sum_{x \in \mathcal{X}} |X - m| \leq D^{\max}$ lies on a regular polytope centered in $[m, ..., m]$ with 2^n faces in an n-dimensional space [1].

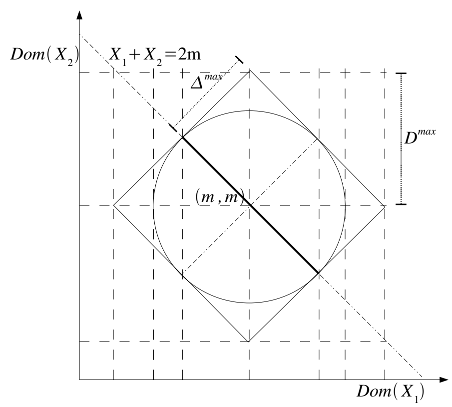

Fig. 3. Relation between SPREAD and DEVIATION for two variables.

The idea is to relax SPREAD with deviation by finding the smallest $D^{\max}$ as possible such that the hyper-sphere is included in the polytope. For two variables X_1 and X_2, Figure 3 shows that the circle can be subsumed by the tangent outer-square. For $D^{\max} = \sqrt{2}\Delta$ the outer-square is tangent with the circle (Pythagore relation). For n variables the result is described in the following theorem.

Theorem 7. *SPREAD*$(\mathcal{X}, m, [0, (\Delta^{\max})^2]) \subseteq$ *DEVIATION*$(\mathcal{X}, m, [0, \sqrt{n}.\Delta^{\max}])$ *and* $\nexists\ (D^{\max} < \sqrt{n}.\Delta)$ *such that SPREAD*$(\mathcal{X}, m, [0, (\Delta^{\max})^2]) \subseteq$ *DEVIATION*$(\mathcal{X}, m, [0, D^{\max}])$.

Proof. For simplicity we assume $m = 0$. Recall that, the set of points such that $\sum_{x \in \mathcal{X}} |X| \leq D^{\max}$ define a regular polytope centered in the origin with 2^n faces in an n-dimensional space. To find $D^{\max}$ such that the polyhedron is tangent with the hyper-sphere of radius Δ, it is easier to work in the positive orthant since others are symmetrical. In this orthant the problem is reduced to finding $D^{\max}$ such that the hyper-plan $X_1 + X_2 + ... + X_n = D^{\max}$ is tangent with the hyper-sphere $X_1^2 + X_2^2 + ... + X_n^2 = (\Delta^{\max})^2$. At the tangent point we have $X_1 =$

[1] For $n = 2$ it is a square and for $n = 3$ it is an octahedron.

$X_2 = ... = X_n$. Consequently at the tangent point $X_1 = X_2 = ... = X_n = \frac{\Delta^{\max}}{\sqrt{n}}$ and hence $D^{\max} = \frac{n}{\sqrt{n}}\Delta^{\max}$. ■

Note that the equality $\texttt{SPREAD}(\mathcal{X}, m, [0, (\Delta^{\max})^2]) = \texttt{DEVIATION}(\mathcal{X}, m, [0, \sqrt{n}.\Delta])$ is valid only when $n = 2$ (two variables). For three variables or more the strict inclusion holds. For instance, the tuple $t = \langle X_1 = \frac{\sqrt{3}}{2}\Delta^{\max}, X_2 = -\frac{\sqrt{3}}{2}\Delta^{\max}, X_3 = 0\rangle \in \texttt{DEVIATION}(\mathcal{X}, m = 0, [0, \sqrt{3}.\Delta^{\max}])$ but $t \notin \texttt{SPREAD}(\mathcal{X}, m = 0, [0, (\Delta^{\max})^2])$. Indeed, $\left(\frac{\sqrt{3}}{2}\Delta^{\max}\right)^2 + \left(\frac{\sqrt{3}}{2}\Delta^{\max}\right)^2 + 0^2 = \frac{3}{2}(\Delta^{\max})^2 > (\Delta^{\max})^2$.

8 Experimental Results

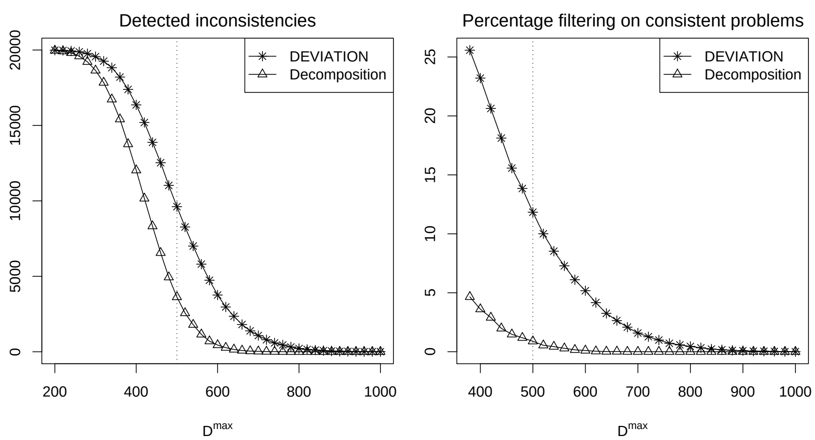

Fig. 4. Experimental Results: DEVIATION v.s. decomposition of Section 3.

The goal of this experiment is to compare the filtering of the DEVIATION propagators described in Section 4 and 5, with an implementation by decomposition as suggested in Section 3.

$20,000$ sets $\mathcal{X} = \{X_1, ..., X_{50}\}$ of random instances were generated. The domain of one variable X are all the integer values between the minimum and maximum of a generated pair of two uniform random integer values between -50 and 50. The mean constraint on each instance is $m = 0.5$. The maximum sum of deviations $D^{\max}$ varies between 200 and 1000. This interval was found experimentally such that for $D^{\max} = 200$ (resp. 1000) all instances are inconsistent (resp. consistent).

The number of inconsistent instances detected by both approaches are given on the left of Figure 4. Note that if an example is detected as inconsistent by decomposition, DEVIATION also detects it. The average percentage of filtering (*i.e.* the number of filtered values divided by the number of initial values in the domains) on consistent instances are plotted on the right of Figure 4.

The number of inconsistent instances detected is significantly larger with the presented propagator than with an implementation by decomposition. For instance, with $D^{\max} = 500$, DEVIATION detects $9,619$ inconsistencies against $3,628$ with the decomposition. On the $10,381$ consistent instances, the pruning percentage obtained with DEVIATION is 11.8% against 0.9% with decomposition.

As shown in Section 6, all inconsistencies are not detected by our propagator if the mean is not integer. Since $m = 0.5$, some inconsistent instances can be undetected. Figure 5 shows a plot of the percentage of inconsistencies detected by decomposition and DEVIATION approaches. Almost all inconsistencies are detected by DEVIATION. The lowest percentage is 99.66% for $D^{\max} = 400$.

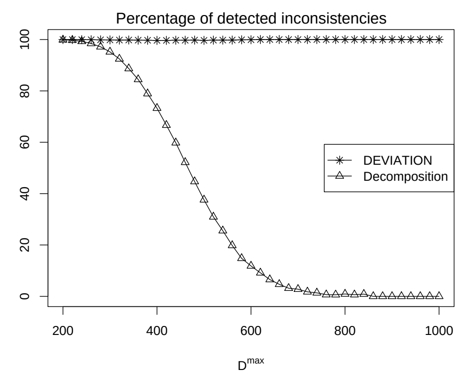

Fig. 5. Percentage of detected inconsistencies.

Section 7 introduces an approximation of SPREAD with DEVIATION. The left of Figure 6 shows the number of inconsistencies detected by SPREAD (as implemented in [7]) compared with the number of inconsistencies found by the approximation. Many inconsistencies remain undetected but, as shown on the left of Figure 6, the propagation using DEVIATION is two orders of magnitude faster than with SPREAD on these 20.000 random instances.

9 Conclusion

This work presents DEVIATION, a new global constraint to balance a set of variables. This constraint is closely related to the SPREAD constraint [5, 7]. While SPREAD constrains the L_2 norm to the mean, DEVIATION constrains the L_1 norm.

The filtering algorithms we introduce run in linear time with respect to the number of variables. Experiments evaluate the efficiency in terms of filtering of our propagators. A relaxation of SPREAD with DEVIATION is also introduced. Such a relaxation offers less filtering but significantly reduces the computation time.

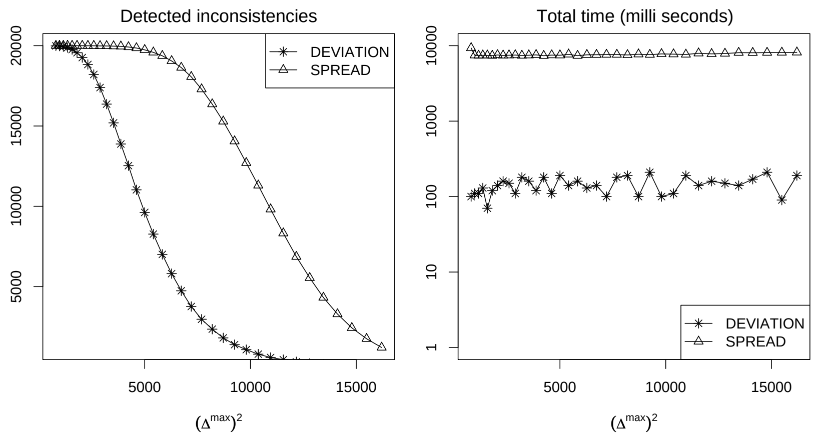

Fig. 6. Experimental Results: `SPREAD` v.s. approximation with `DEVIATION`.

Acknowledgment

The authors wish to thank Julien Hendrickx and Raphaël Jungers for their proof of Theorem 5 and also Jean-Noël Monette and Grégoire Dooms for various interesting discussions.

This research is supported by the Walloon Region, project Transmaze (516207).

References

1. *Problem 30 of CSPLIB (www.csplib.org).*
2. C. Castro and S. Manzano. Variable and value ordering when solving balanced academic curriculum problem. *Proc. of the ERCIM WG on constraints*, 2001.
3. M. R. Garey and David S. Johnson. *Computer and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.
4. Brahim Hnich, Zeynep Kiziltan, and Toby Walsh. Modelling a balanced academic curriculum problem. *Proceedings of CP-AI-OR-2002*, 2002.
5. G. Pesant and J.C. Régin. Spread: A balancing constraint based on statistics. *Lecture Notes in Computer Science*, 3709:460–474, 2005.
6. J.C. Régin, T. Petit, C. Bessiere, and J.-F. Puget. An original constraint based approach for solving over constrained problems. *Sixth International Conference on Principles and Practice of Constraint Programming (CP 2000)*, 1894, 2000.
7. P. Schaus, Y. Deville, P. Dupont, and J.C. Régin. Simplication and extension of the spread constraint. *Third International Workshop on Constraint Propagation And Implementation*, 2006.
8. W.J. van Hoeve, G. Pesant, L.M. Rousseau, and A. Sabharwal. Revisiting the sequence constraint. *Twelfth International Conference on Principles and Practice of Constraint Programming (CP 2006)*, 4204, 2006.