

La Contrainte Déviation *

Pierre Schaus¹, Yves Deville¹, Pierre Dupont¹, Jean-Charles Régin²

¹Université catholique de Louvain, Belgium

²ILOG, France

¹{pschaus, yde, pdupont}@info.ucl.ac.be, ²regin@ilog.fr

Résumé

Cet article présente DEVIATION, une contrainte globale afin d'obtenir des solutions équilibrées. La norme L_p du vecteur des variables diminuées de leur moyenne est une mesure de violation de l'équilibre parfait. SPREAD contraint la somme du carré des écarts à la moyenne, c'est-à-dire la norme L_2 . Cet article considère la norme L_1 . Aucun de ces deux critères n'est plus général que l'autre. Néanmoins, la conception des propagateurs pour L_1 est plus simple. De plus la complexité temporelle de ceux-ci est linéaire par rapport au nombre de variables alors qu'elle est quadratique pour les propagateurs de SPREAD.

1 Introduction

Nous utilisons le problème d'équilibrage des horaires académiques [1] (BACP ¹) pour illustrer l'utilité de DEVIATION. L'objectif est d'assigner une période à chaque cours tout en respectant les pré-requis entre les cours et en équilibrant autant que possible les charges de travail entre les périodes. Le BACP est donc un problème d'optimisation avec pour objectif de maximiser l'équilibre.

Dans le BACP, la charge moyenne m des périodes est connue puisque le nombre de crédits de chaque cours est fixé. Il n'existe généralement pas de solution parfaitement équilibrée c'est-à-dire ayant pour chaque période une charge de m . Néanmoins, il est possible de définir une mesure de violation de l'équilibre parfait. Pour un ensemble de variables $\mathcal{X} = \{X_1, X_2, \dots, X_n\}$ et une moyenne fixée m , une mesure de violation de l'équilibre parfait peut-être définie comme la norme L_p du vecteur $[\mathbf{X} - \mathbf{m}]$ avec $\mathbf{X} = [X_1, X_2, \dots, X_n]$, $\mathbf{m} = [m, m, \dots, m]$ tel que $\sum_{i=1}^n X_i = n.m$. La norme L_p de

$[\mathbf{X} - \mathbf{m}]$ est définie comme $(\sum_{i=1}^n |X_i - m|^p)^{\frac{1}{p}}$ avec $p \geq 0$.

Suivant l'idée proposée par Régin et al. [6] pour lâcher les contraintes globales, nous définissons une violation de l'équilibre parfait comme une variable de coût L_p dans la contrainte globale d'équilibre : **soft-balance** (\mathcal{X}, m, L_p) est satisfaite si et seulement si la norme L_p du vecteur $[\mathbf{X} - \mathbf{m}]$ est égale à L_p et $\sum_{i=1}^n X_i = n.m$.

Une interprétation de la mesure de violation de l'équilibre pour quelques normes spécifiques est donnée ci-dessous :

- L_0 : $|\{X \in \mathcal{X} | X \neq m\}|$ est le nombre de valeurs différentes de la moyenne m .
- L_1 : $\sum_{X \in \mathcal{X}} |X - m|$ est la somme des écarts à la moyenne.
- L_2 : $\sum_{X \in \mathcal{X}} (X - m)^2$ est la somme des écarts quadratiques à la moyenne.
- L_∞ : $\max_{X \in \mathcal{X}} |X - m|$ est le plus grand écart à la moyenne.

Notons qu'aucun des critères énoncés ci-dessus n'est plus général qu'un autre. Par exemple, une minimisation de L_1 n'implique pas nécessairement une minimisation du critère L_2 . Cela est illustré sur l'exemple suivant. Supposons qu'un problème sous contraintes admette quatre solutions données dans le Tableau 1. La solution la plus équilibrée dépend de la norme considérée. En effet, chaque solution présente une moyenne de 100 mais est optimale selon une norme différente.

La norme L_∞ a déjà été utilisée dans deux travaux [2, 4] pour résoudre le BACP. SPREAD est une contrainte globale pour la norme L_2 [5, 7]. Une contrainte pour L_0 peut être aisément implémentée à l'aide d'une contrainte **AT-LEAST** $(i, [X_1, \dots, X_n], m)$ pour $|\{X \in \mathcal{X} | X \neq m\}| \leq i$ et une contrainte **SUM** $([X_1, \dots, X_n], n.m)$ pour imposer

*Traduction française de l'article [8] présenté à CPAIOR 07
¹pour Balanced Academic Curriculum Problem

	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

TAB. 1 – Illustration montrant qu’aucun critère défini par les normes L_0 , L_1 , L_2 or L_∞ n’est plus général qu’un autre. La plus petite norme est indiquée en gras. Par exemple, la solution 2 est la plus équilibrée selon le critère L_1 .

une moyenne m . Cet article propose une contrainte globale et ses propagateurs pour la norme L_1 . Cette contrainte est formulée dans la définition suivante :

Définition 1. *Étant donné un ensemble de variables à domaines finis $\mathcal{X} = \{X_1, X_2, \dots, X_n\}$, une valeur moyenne m et une variable intervalle D , la contrainte $DEVIATION(\mathcal{X}, m, D)$ spécifie que la moyenne des valeurs prises par les variables de \mathcal{X} est égale à m et que la somme des écarts à m est égale à D . Plus exactement, $DEVIATION(\mathcal{X}, m, D)$ est satisfaite si et seulement si*

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

Pour que la contrainte soit consistante, $n.m$ doit être entier. Par conséquent $n.D$ est également entier.

Plan de l’article :

La Section 2 introduit des notions relatives à la programmation par contraintes telles que la propagation, la consistance de domaine et la consistance aux bornes. Nous y définissons également quelques notations utiles pour la suite. La Section 3 donne les motivations pour la contrainte globale $DEVIATION$ en termes de propagation. La section 4 explique les propagateurs réduisant le domaine de la variable D . Ces propagateurs utilisent la minimisation et maximisation de la somme des écarts à la moyenne. La minimisation peut être résolue en temps linéaire. La maximisation est un problème \mathcal{NP} -complet mais une borne supérieure peut être calculée en temps linéaire. La Section 5 décrit un algorithme de filtrage depuis m et D vers les variables de \mathcal{X} . L’idée est similaire à un algorithme de consistance aux bornes pour la contrainte SUM mais incluant la contrainte de la somme des écarts à la moyenne. La Section 6 montre que nos propagateurs n’atteignent pas toujours la consistance aux bornes en particulier lorsque la moyenne n’est pas entière. La Section 7 donne une relaxation de $SPREAD$ à l’aide de $DEVIATION$.

Pour terminer, la Section 8 évalue les performances des propagateurs présentés en termes de filtrage sur des instances aléatoires.

2 Notions préliminaires et notations

Quelques notions de programmation par contraintes (inspirées de la Section 2 de [9]) sont introduites. Ensuite, quelques notations utiles pour le reste de l’article sont définies.

Soit X une variable à domaine fini entier. Son *domaine* est un ensemble ordonné d’entiers pouvant être assigné à X et dénoté $Dom(X)$. La valeur minimum (resp. maximum) de son domaine est dénotée $X^{\min} = \min(Dom(X))$ (resp. $X^{\max} = \max(Dom(X))$). Soit $\mathcal{X} = \{X_1, X_2, \dots, X_k\}$ une séquence de variables. Une *contrainte* C sur \mathcal{X} est définie comme un sous-ensemble du produit cartésien des domaines des variables dans \mathcal{X} : $C \subseteq Dom(X_1) \times Dom(X_2) \times \dots \times Dom(X_k)$. Un nuplet $(v_1, \dots, v_k) \in C$ est appelé une *solution* de C . Une valeur $v \in Dom(X_i)$ pour $i \in 1, \dots, k$ est *inconsistante* par rapport à C si elle n’appartient à aucun nuplet de C , sinon elle est *consistante*. La contrainte C est inconsistante si elle ne contient aucune solution. Sinon, C est dite consistante.

Le processus de recherche de solution(s) en programmation par contraintes alterne la *propagation* des contraintes et la recherche. La recherche consiste en l’énumération de toutes les combinaisons possibles variables-valeurs jusqu’à ce qu’une solution soit trouvée où qu’il soit prouvé qu’il n’en existe pas. Nous disons que ce processus construit un arbre de recherche. Afin de réduire le nombre exponentiel de combinaisons, la propagation des contraintes est appliquée à chaque noeud de l’arbre de recherche : Pour une contrainte C et des domaines pour les variables impliquées dans C , un propagateur pour C retire des domaines des valeurs n’appartenant à aucune solution de C . Cela est répété pour chaque contrainte jusqu’à ce qu’aucune valeur ne puisse être retirée par les propagateurs. La suppression de valeurs inconsistantes dans les domaines est aussi appelé *filtrage*.

Pour que la recherche soit efficace, les algorithmes de filtrage doivent être rapides car ils sont exécutés de nombreuses fois durant le processus de recherche. Ils doivent également retirer le plus possible de valeurs inconsistantes. Si un algorithme de filtrage pour une contrainte C retire toute les valeurs inconsistantes des domaines par rapport à C , nous disons que l’algorithme atteint la consistance de domaines pour C . Pour certaines contraintes telles que $AllDiff$, il est possible d’atteindre la consistance de domaines en temps polynomial mais pour d’autres telles que SUM , cela serait bien trop coûteux. Dans de tels cas, il de-

vient avantageux d'utiliser une notion de consistance plus faible par exemple la consistance aux bornes. Une contrainte C est bornes-consistante si les bornes des domaines de chaque variable impliquée dans C appartiennent à au moins à une solution de C . L'idée est de borner le domaine de chaque variable par un intervalle en s'assurant que les points extrêmes des intervalles satisfasse la propriété de consistance de domaines. Si ce n'est pas le cas, les bornes supérieures et inférieures des intervalles peuvent être ajustées jusqu'à ce que la consistance de bornes soit obtenue.

La Proposition 1 affirme que réaliser la consistance de domaines pour DEVIATION est aussi difficile que pour les contraintes arithmétiques en général.

Proposition 1. *Atteindre la consistance de domaine pour DEVIATION est \mathcal{NP} -Complet.*

Démonstration. La contrainte $\text{SUM}(X_1 \dots X_n, S)$ est satisfaite si et seulement si $\sum_{i=1}^n X_i = S$. Il est bien connu que réaliser la consistance de domaines pour SUM est un problème \mathcal{NP} -Complet (Le problème *subset sum* [3] peut aisément être réduit vers le problème de réalisation de la consistance de domaines pour SUM). Dans le cas particulier où $\text{Dom}(D) = [0, +\infty]$ et $m = S/n$, réaliser la consistance de domaine pour SUM se réduit à réaliser la consistance de domaines pour DEVIATION(\mathcal{X}, m, D). \square

Dans la suite, les domaines des variables sont considérés comme pleins et décrits par l'intervalle $\text{Dom}(X) = [X^{\min} \dots X^{\max}]$.

La Définition 2 introduit quelques notations utiles qui peuvent être aisément comprises à l'aide de la Figure 1. Un exemple numérique est également donné à l'Exemple 1.

Définition 2. *Pour une variables X et une valeur donnée m , les bornes supérieures sur l'écart à droite et à gauche de la moyenne sont respectivement*

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

Les sommes de ces valeurs sur l'ensemble \mathcal{X} sont respectivement

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

Suivant la même idée, les bornes inférieures sur les écarts à la moyenne sont :

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

Pour une variable $X_i \in \mathcal{X}$ nous définissons :

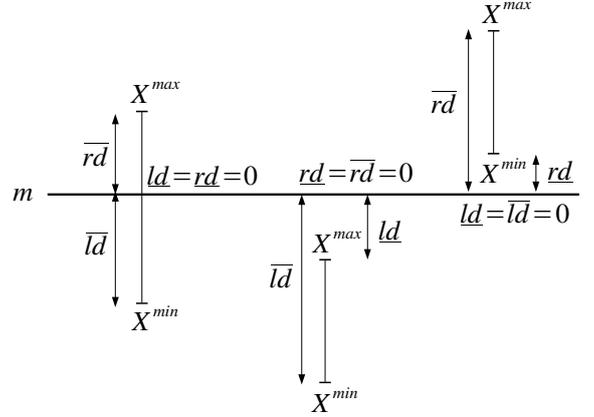


FIG. 1 – Illustration des définitions introduites à la Définition 2 pour trois variables.

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

Afin d'alléger les notations, l'argument (\mathcal{X}, m) sera parfois omis. Par exemple $\overline{LD}(\mathcal{X}, m)$ s'écrit simplement \overline{LD} .

Exemple 1. *Soit quatre variables $\mathcal{X} = \{X_1, X_2, X_3, X_4\}$ ayant pour domaines $\text{Dom}(X_1) = [8, 10]$, $\text{Dom}(X_2) = [4, 7]$, $\text{Dom}(X_3) = [1, 5]$ et $\text{Dom}(X_4) = [3, 4]$. Le tableau suivant donne les quantités introduites à la Définition 2.*

i	X_i^{\min}	X_i^{\max}	$\overline{rd}(X_i, 5)$	$\overline{ld}(X_i, 5)$	$\underline{rd}(X_i, 5)$	$\underline{ld}(X_i, 5)$
1	8	10	5	0	3	0
2	4	7	2	1	0	0
3	1	5	0	4	0	0
4	3	4	0	2	0	1
\sum_i			7	7	3	1

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

Le filtrage pour DEVIATION est basé sur le théorème suivant stipulant que les sommes des écarts au dessus et en dessous de la moyenne doivent être égales.

Lemme 1. *Soit $\mathcal{X} = \{X_1, \dots, X_n\}$. L'égalité $n.m = \sum_{X \in \mathcal{X}} X$ est satisfaite si et seulement si $\sum_{X > m} (X - m) = \sum_{X < m} (m - X)$.*

Démonstration. $n.m = \sum_{X \in \mathcal{X}} X$ peut se reformuler comme $0 = \sum_{X \in \mathcal{X}} X - n.m = \sum_{X > m} (X - m) + \sum_{X < m} (X - m) + \sum_{X = m} (X - m) = \sum_{X > m} (X - m) - \sum_{X < m} (m - X)$. \square

Propriété 1. Soit $\mathcal{X} = \{X_1, \dots, X_n\}$. Une affectation sur \mathcal{X} satisfait :

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

Théorème 1. $DEVIATION(\mathcal{X}, m, D)$ est consistante seulement si les conditions suivantes sont satisfaites :

1. $\underline{RD} \leq \frac{D^{\max}}{2}$
2. $\underline{LD} \leq \frac{D^{\max}}{2}$
3. $\overline{RD} \geq \frac{D^{\min}}{2}$
4. $\overline{LD} \geq \frac{D^{\min}}{2}$
5. $[\underline{LD}, \overline{LD}] \cap [\underline{RD}, \overline{RD}] \neq \emptyset$

Démonstration.

1. Si $\underline{RD}(\mathcal{X}, m) > \frac{D^{\max}}{2}$ alors $\sum_{X > m} (X - m) > \frac{D^{\max}}{2}$ (Propriété 1). Donc $\sum_{i=1}^n |X_i - m| > D^{\max}$ (par le Lemme 1).
- 2., 3. et 4. sont similaires à 1.
5. est une conséquence des Lemme 1 et Propriété 1. \square

3 Implémentation Naïve

Cette section explique pourquoi une implémentation naïve de $DEVIATION$ par décomposition en plusieurs contraintes élémentaires n'est pas optimale en termes de filtrage.

Comme il est dit dans la Définition 1, $DEVIATION(\mathcal{X}, m, D)$ est satisfaite si et seulement si $\sum_{i=1}^n X_i = n.m$ et $D = \sum_{i=1}^n |X_i - m|$. Cela suggère une implémentation de la contrainte par décomposition en deux contraintes SUM. La Figure 2 montre que le filtrage obtenu avec la décomposition n'est pas optimal.

Soit deux variables X_1, X_2 avec des domaines non bornés et la contrainte

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

Le carré hachuré de lignes diagonales (voire Figure 2) délimite l'ensemble des points tels que $|X_1 - m| + |X_2 - m| \leq D^{\max}$. La ligne diagonale est l'ensemble des points tels que $X_1 + X_2 = 2m$. Les domaines non bornés pour X_1 et X_2 sont bornes-consistants pour la contrainte moyenne. Le rectangle hachuré verticalement définit le domaine de X_1 après un filtrage consistant aux bornes pour $|X_1 - m| + |X_2 - m| \leq D^{\max}$. L'ensemble des solutions de $DEVIATION$ est défini par le segment diagonal en gras obtenu par intersection de la surface carrée et de la ligne diagonale. La figure montre que les domaines peuvent être davantage réduits. En effet, le filtrage bornes-consistant sur le segment diagonal en gras donne un domaine pour X_1 représenté par

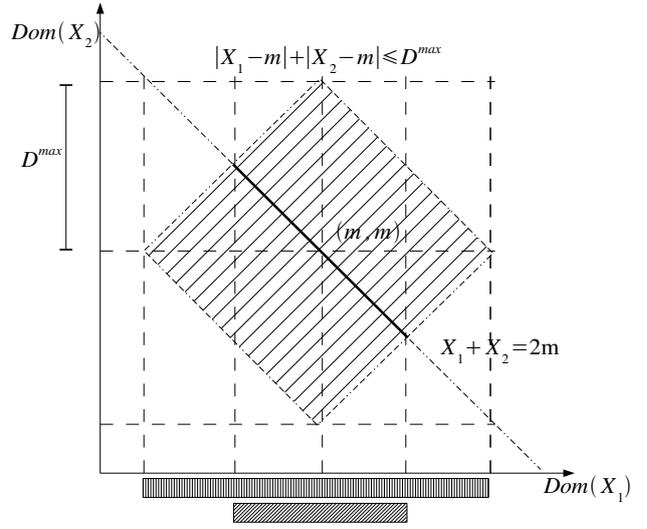


FIG. 2 – Filtrage de X_1 avec la décomposition et avec $DEVIATION$.

le rectangle hachuré de lignes diagonales. En conclusion, un filtrage borne-consistant pour la décomposition donne $Dom(X_1) = Dom(X_2) = [m - D^{\max}, m + D^{\max}]$ alors qu'un filtrage consistant aux bornes pour $DEVIATION(\{X_1, X_2\}, m, D \in [0, D^{\max}])$ donne des domaines deux fois plus petits $Dom(X_1) = Dom(X_2) = [m - \frac{D^{\max}}{2}, m + \frac{D^{\max}}{2}]$.

4 Filtrage de D

Le filtrage bornes-consistant de D nécessite de résoudre deux problèmes d'optimisation : la minimisation et la maximisation de la somme des écarts à la moyenne donnée m . La Définition 3 définit une relaxation de ces problèmes en permettant une assignation sur les rationnels.

Définition 3. \underline{D} et \overline{D} dénotent les valeurs optimales des problèmes :

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

tels que :

$$\begin{aligned} \sum_{i=1}^n X_i &= n.m \\ X_i^{\min} &\leq X_i \leq X_i^{\max}, \quad 1 \leq i \leq n \\ X_i &\in \mathbb{Q}, \quad 1 \leq i \leq n. \end{aligned}$$

Les valeurs \underline{D} et \overline{D} peuvent être utilisées pour filtrer le domaine de D :

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

Le restant de la section explique comment \underline{D} peut être calculé en temps linéaire par rapport au nombre de variables $n = |\mathcal{X}|$. Malheureusement, trouver \underline{D} est un problème \mathcal{NP} -complet et le mieux que nous puissions faire est de créer une bonne borne supérieure comme expliqué en fin de section.

La Définition 4 caractérise une solution optimale au problème de trouver \underline{D} .

Définition 4 (affectation centrée vers le haut et vers le bas). Soit $\mathcal{X} = \{X_1, \dots, X_n\}$. Soit $A : X_i \rightarrow \text{Dom}(X_i)$ une affectation de $X \in \mathcal{X}$. La quantité $s(A)$ dénote la somme des valeurs affectées : $s(A) = \sum_{X_i \in \mathcal{X}} A(X_i)$.

Une affectation A est dit **centrée vers le haut** quand :

$$A(X_i) \begin{cases} = X_i^{\min} & \text{si } X_i^{\min} \geq s(A)/n \\ \leq s(A)/n & \text{sinon} \end{cases}$$

En d'autres termes, chaque variable dont la plus petite valeur de son domaine est supérieure à la moyenne des valeurs affectées prend la valeur minimum de son domaine et les autres variables prennent des valeurs inférieures à la moyenne des valeurs affectées.

Un affectation A est dit **centrée vers le bas** quand :

$$A(X_i) \begin{cases} = X_i^{\max} & \text{si } X_i^{\max} \leq s(A)/n \\ \geq s(A)/n & \text{sinon} \end{cases}$$

En d'autres termes, chaque variable dont la plus grande valeur de son domaine est inférieure à la moyenne des valeurs affectées prend la valeur maximum de son domaine et les autres variables prennent des valeurs supérieures à la moyenne des valeurs affectées.

Exemple 2. Considérant les variables et domaines de l'Exemple 1, l'affectation suivante est centrée vers le haut avec une moyenne de $17/4$:

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

Théorème 2. Une affectation est une solution optimale au problème de trouver \underline{D} si et seulement si elle est centrée vers le haut ou vers le bas avec une moyenne m .

Démonstration. (si) Étant donné un affectation A de moyenne m i.e. $s(A) = n.m$. La seule manière de diminuer la somme des écarts tout en conservant une moyenne de m est de trouver une paire de variables (X_i, X_j) telle que $A(X_i) > m, A(X_i) > X_i^{\min}, A(X_j) < m, A(X_j) < X_j^{\max}$ et de les rapprocher toutes les deux de m en diminuant $A(X_i)$ et augmentant $A(X_j)$ d'une même quantité. Par définition d'une affectation centrée vers le haut et

vers le bas, il est impossible de trouver une telle paire de variables X_i, X_j . Par conséquent, les affectations centrées vers le haut ou vers le bas sont des solutions optimales. (seulement si) Supposons une affectation A qui n'est ni centrée vers le haut, ni centrée vers le bas telle que $s(A) = n.m$. Il est alors possible de trouver au moins deux variables $X_i, X_j \in \mathcal{X}$. La première telle que $A(X_i) > m$ et $A(X_i) > X_i^{\min}$ (violation de centrée vers le haut) et la seconde telle que $A(X_j) < m$ et $A(X_j) < X_j^{\max}$ (violation de centrée vers le bas). Nous définissons $\delta = \min(A(X_i) - \max(X_i^{\min}, m), \min(X_j^{\max}, m) - A(X_j))$. L'affectation $A(X)$ n'est pas optimale puisque la somme des écarts à m peut-être diminuée de 2δ en modifiant l'affectation sur X_i et X_j : $A'(X_i) = A(X_i) - \delta$ et $A'(X_j) = A(X_j) + \delta$. \square

Théorème 3. Si *DEVIATION* est consistante alors

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

Démonstration. Supposons $\underline{LD} \geq \underline{RD}$. Il est alors possible de construire un affectation A centrée vers le bas de moyenne m qui est optimal (par le Théorème 2). Pour cette affectation $\sum_{A(X) < m} (m - A(X)) = \underline{LD}$ (par la Définition 2 de \underline{LD}). Étant donné que $\sum_{X > m} (X - m) = \sum_{X < m} (m - X)$ (par le Lemme 1), la somme des écarts de cette affectation centrée vers le bas est $\sum_{X \in \mathcal{X}} |A(X) - m| = 2.\underline{LD}$.

Le cas $\underline{LD} \leq \underline{RD}$ est similaire. L'affectation est centrée vers le haut plutôt que vers le bas. \square

Exemple 3. Les variables et domaines sont identiques à ceux de l'Exemple 1. La moyenne considérée est $m = 5$. En utilisant les valeurs calculées $\underline{LD}(\mathcal{X}, 5) = 1$ et $\underline{RD}(\mathcal{X}, 5) = 3$ dans l'Exemple 1, il peut être déduit que $\underline{D} = 2. \max(1, 3) = 6$. Par conséquent, le filtrage sur D pour *DEVIATION* ($\mathcal{X} = \{X_1, X_2, X_3, X_4\}, m = 5, D \in [0, 7]$) conduit à $\text{Dom}(D) = [6, 7]$.

Théorème 4. Calculer \bar{D} est un problème \mathcal{NP} -complet.

Démonstration. Ce problème est aussi difficile que dans le cas particulier où $m = 0$:

$$\bar{D} = \max \sum_{i=1}^n |X_i|$$

tel que :

$$\sum_{i=1}^n X_i = 0$$

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

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

Il est possible de réduire le problème *subset sum* [3] vers le problème de trouver \bar{D} . Étant donné un ensemble de n valeurs positives $\{b_1, \dots, b_{n-1}, T\}$, le problème *subset sum* est de trouver s'il existe un ensemble de valeurs binaires $\{y_1, \dots, y_{n-1}\}, y_i \in \{0, 1\}, 1 \leq i < n$ telles que $\sum_{i=1}^{n-1} y_i \cdot b_i = T$. La réduction est la suivante :

- $X_i^{\min} = -\frac{b_i}{2}$ et $X_i^{\max} = \frac{b_i}{2}$ pour $1 \leq i < n$.
- $X_n = \frac{\sum_{i=1}^{n-1} b_i}{2} - T$.
- Il y a une solution au problème *subset sum* si et seulement si $\bar{D} \geq \frac{\sum_{i=1}^{n-1} b_i}{2} + \left| \frac{\sum_{i=1}^{n-1} b_i}{2} - T \right|$. Cette contrainte sur la valeur optimale garantit que la solution optimale est telle que $X_i \in \{-\frac{b_i}{2}, \frac{b_i}{2}\}$. La solution au problème *subset sum* est alors donnée par $y_i = 1$ si $X_i = \frac{b_i}{2}$ et $y_i = 0$ si $X_i = -\frac{b_i}{2}$.

□

A moins que $\mathcal{P} = \mathcal{NP}$, le problème de calculer \bar{D} est exponentiel (Théorème 4). Comme expliqué à la Section 2, un algorithme de filtrage doit être aussi rapide que possible pour être utile. En effet celui-ci est exécuté de nombreuses fois durant le processus de recherche. C'est pour cette raison que nous préférons trouver rapidement une borne supérieure \bar{D}^\dagger pour \bar{D} plutôt que sa valeur exacte. Une borne supérieure pouvant être calculée en temps linéaire est

$$\bar{D}^\dagger = \sum_{i=1}^n \max(|X_i^{\max} - m|, |X_i^{\min} - m|).$$

Le filtrage de $Dom(D)$ devient :

$$Dom(D) \longleftarrow Dom(D) \cap [\underline{D}, \bar{D}^\dagger]$$

5 Filtrage de \mathcal{X}

Deux propagateurs peuvent être considérés pour filtrer le domaine de \mathcal{X} :

1. depuis D^{\min} et m vers \mathcal{X} et
2. depuis D^{\max} et m vers \mathcal{X} .

Atteindre la consistance aux bornes pour le premier propagateur est \mathcal{NP} -complet. En effet, vérifier la consistance d'une valeur nécessite de maximiser la somme des écarts ce qui est un problème \mathcal{NP} -complet (Théorème 4). La décomposition de DEVIATION présentée dans la Section 3 peut néanmoins être utilisée pour réaliser la consistance de bornes pour la contrainte $\sum_{X \in \mathcal{X}} |X - m| \geq D^{\min}$. Remarquons que ce filtrage est inutile si nous cherchons une solution équilibrée sur \mathcal{X} . Le restant de cette section se concentre donc sur un algorithme pour le second propagateur.

Le filtrage est basé sur le calcul des valeurs \bar{X}_i et \underline{X}_i introduites dans la Définition 5.

Définition 5. \bar{X}_i et \underline{X}_i sont les valeurs optimales des problèmes suivant :

$$\bar{X}_i = \max(X_i) \quad \text{et} \quad \underline{X}_i = \min(X_i)$$

tels que :

$$\sum_{j=1}^n X_j = n \cdot m \tag{1}$$

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

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

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

La règle de filtrage du domaine de X_i s'écrit simplement :

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

Théorème 5. Pour une variable X_i , si la contrainte est consistante, les égalités suivantes sont satisfaites :

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

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

Démonstration. Seul \bar{X}_i est étudié car la preuve pour \underline{X}_i est symétrique par rapport à m . Deux cas sont possibles :

- $\bar{LD}_i \leq \frac{D^{\max}}{2}$: Par le Lemme 1 les sommes des écarts au dessus et en dessous de la moyenne doivent être égales. Donc la solution optimale est telle que $\bar{X}_i - m + \underline{RD}_i = \bar{LD}_i$. La contrainte (2) n'est pas serrée² dans ce cas.
- $\bar{LD}_i > \frac{D^{\max}}{2}$: Par le Lemme 1, la contrainte (1) signifie que la somme des écarts au dessus et en dessous de la moyenne doivent être égales. La conjonction des contraintes (1) et (2) signifie que les sommes des écarts au dessus et en dessous de la moyenne sont égales et valent au plus $D^{\max}/2$. Par conséquent la solution optimale est telle que $\bar{X}_i - m + \underline{RD}_i = \frac{D^{\max}}{2}$. La contrainte (2) est serrée dans ce cas.

Si les deux cas sont considérés ensemble, l'égalité suivante est satisfaite pour une solution optimale $\bar{X}_i - m + \underline{RD}_i = \min(\frac{D^{\max}}{2}, \bar{LD}_i)$. □

La procédure de filtrage sur \mathcal{X} applique la règle (3) une fois sur chaque $X_i \in \mathcal{X}$. Cela se fait en temps linéaire par rapport au nombre de variables.

²Une contrainte d'inégalité est dite non serrée si elle est vérifiée à l'aide du signe d'inégalité stricte et serrée si elle est vérifiée à l'aide du signe d'égalité

Exemple 4. Les variables et domaines sont les mêmes que dans l'Exemple 1. La contrainte étudiée est $DEVIATION(\mathcal{X} = \{X_1, X_2, X_3, X_4\}, m = 5, D \in [0, 7])$. Les valeurs \bar{X}_i et \underline{X}_i sont : $\bar{X}_1 = \min(3.5, 7) - 0 + 5 = 8.5$, $\bar{X}_2 = \min(3.5, 6) - 3 + 5 = 5.5$, $\bar{X}_3 = \min(3.5, 3) - 3 + 5 = 5$, $\bar{X}_4 = \min(3.5, 5) - 3 + 5 = 5.5$, $\underline{X}_1 = -\min(3.5, 2) + 1 + 5 = 4$, $\underline{X}_2 = -\min(3.5, 5) + 1 + 5 = 2.5$, $\underline{X}_3 = -\min(3.5, 7) + 1 + 5 = 2.5$ et $\underline{X}_4 = -\min(3.5, 7) + 0 + 5 = 1.5$. La règle (3) donne les domaines filtrés : $Dom(X_1) = [8, 8]$, $Dom(X_2) = [4, 5]$, $Dom(X_3) = [3, 5]$ et $Dom(X_4) = [3, 4]$.

6 La consistance de bornes pour DEVIATION

Le filtrage pour toutes les variables s'effectue en $\mathcal{O}(n)$ comme suit :

1. Filtrage depuis D^{\max} et m vers \mathcal{X} :
 $\forall X \in \mathcal{X}, Dom(X) \leftarrow Dom(X) \cap [\underline{X}, \bar{X}]$.
2. Filtrage depuis \mathcal{X} et m vers D :
 $Dom(D) \leftarrow Dom(D) \cap [\underline{D}, \bar{D}^\uparrow]$.

Même s'il était possible de calculer \bar{D} efficacement, la consistance de bornes ne serait pas nécessairement obtenue. La raison étant que que les valeurs \bar{X} , \underline{X} (Définition 5) et \underline{D} (Définition 3) sont calculées en faisant l'hypothèse que les domaines des variables sont des intervalles rationnels (\mathbb{Q}) plutôt qu'entiers (\mathbb{Z}). Comme le montre l'exemple suivant, cela a pour effet de ne pas filtrer certaines valeurs inconsistantes.

Exemple 5. Considérons une ensemble de 10 variables \mathcal{X} avec pour domaine $[0, 1]$ et une contrainte moyenne $m = 0.5$. Le Théorème 3 donne la valeur $\underline{D} = 0$ puisque chaque domaine chevauche m . En réalité, la seule façon possible d'obtenir une affectation qui respecte la contrainte moyenne est d'assigner cinq variables à 0 et cinq variables à 1. Pour une telle affectation, la somme des écarts à la moyenne minimum vaut 5 et non 0. En conséquence, la contrainte $DEVIATION(\mathcal{X}, m = 0.5, D \in [0, 3])$ est inconsistante. Cette inconsistance n'est pas détectée par notre propagateur.

L'Exemple 5 montre que toutes les inconsistances ne sont pas détectées par notre propagateur. Cela se produit lorsque la moyenne n'est pas entière mais rationnelle et lorsque les domaines de certaines variables chevauchent la moyenne. Lorsque la moyenne est entière, le propagateur est bornes-consistant.

7 Lien entre SPREAD et DEVIATION

Cette section montre que DEVIATION peut être utilisée comme relaxation de SPREAD. Cette relaxation peut

être utile puisque la propagation pour SPREAD s'effectue en $\mathcal{O}(n^2)$ contre $\mathcal{O}(n)$ pour DEVIATION. De plus, DEVIATION est plus simple à implémenter. La relaxation est illustrée sur un schéma et les paramètres à donner à DEVIATION pour obtenir la plus forte relaxation possible sont donnés en fonction des paramètres de SPREAD.

$SPREAD(\mathcal{X}, m, \Delta^2)$ est satisfaite si la valeur m est la moyenne sur \mathcal{X} et que la somme des écarts quadratiques à la moyenne m vaut Δ^2 . Plus formellement, SPREAD est satisfaite si $\sum_{x \in \mathcal{X}} X = n.m$ et $\sum_{x \in \mathcal{X}} (X - m)^2 = \Delta^2$.

D'un point de vue géométrique, $\sum_{x \in \mathcal{X}} (X - m)^2 \leq (\Delta^{\max})^2$ définit une hyper-sphère centrée en $[m, \dots, m]$ de rayon Δ^{\max} . L'ensemble des points satisfaisant $\sum_{x \in \mathcal{X}} |X - m| \leq D^{\max}$ se trouve sur un polytope régulier centré en $[m, \dots, m]$ avec 2^n facettes dans un espace à n dimensions ³.

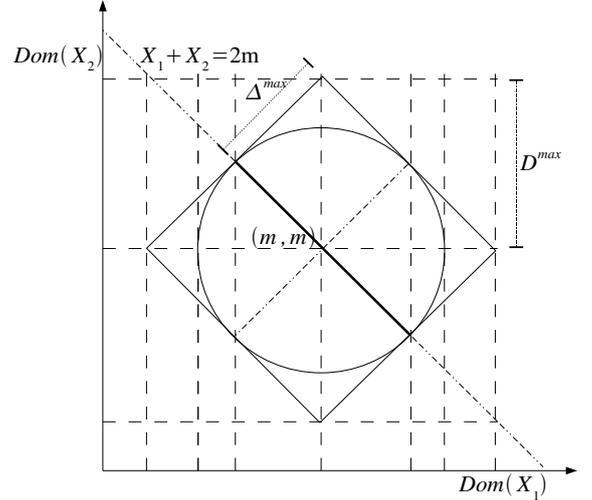


FIG. 3 – Lien entre SPREAD et DEVIATION pour deux variables.

L'idée est de relâcher SPREAD avec DEVIATION en trouvant le plus petit D^{\max} possible tel que l'hyper-sphère est incluse dans le polytope. Pour deux variables X_1 et X_2 , la Figure 3 montre que le cercle peut être placé dans un carré externe tangent. Pour $D^{\max} = \sqrt{2}\Delta$ le carré externe est tangent avec le cercle (relation de Pythagore). Pour n variables le résultats est décrit dans le Théorème suivant.

Théorème 6. $SPREAD(\mathcal{X}, m, [0, (\Delta^{\max})^2]) \subseteq DEVIATION(\mathcal{X}, m, [0, \sqrt{n} \cdot \Delta^{\max}])$ et $\nexists D^{\max} < \sqrt{n} \cdot \Delta^{\max}$ tel que $SPREAD(\mathcal{X}, m, [0, (\Delta^{\max})^2]) \subseteq DEVIATION(\mathcal{X}, m, [0, D^{\max}])$.

³Pour $n = 2$ c'est un carré et pour $n = 3$ c'est un octaèdre.

Démonstration. Pour des raisons de simplicité nous considérons $m = 0$. L'ensemble des points tels que $\sum_{x \in \mathcal{X}} |X| \leq D^{\max}$ définit un polytope régulier centré à l'origine à 2^n facettes dans un espace à n dimensions. Nous travaillons dans le premier orthant pour trouver D^{\max} tel que le polytope est tangent à l'hyper-sphère de rayon Δ (la situation étant symétrique dans les autres orthants). Dans cet orthant, le problème est réduit à trouver D^{\max} tel que l'hyperplan $X_1 + X_2 + \dots + X_n = D^{\max}$ est tangent avec l'hypersphère $X_1^2 + X_2^2 + \dots + X_n^2 = (\Delta^{\max})^2$. Au point de tangence, nous avons $X_1 = X_2 = \dots = X_n$. Par conséquent, en ce point $X_1 = X_2 = \dots = X_n = \frac{\Delta^{\max}}{\sqrt{n}}$ et donc $D^{\max} = \frac{n}{\sqrt{n}} \Delta^{\max}$. \square

Notons que l'égalité $\text{SPREAD}(\mathcal{X}, m, [0, (\Delta^{\max})^2]) = \text{DEVIATION}(\mathcal{X}, m, [0, \sqrt{n} \cdot \Delta])$ n'est valide que quand $n = 2$ (deux variables). Pour trois variables ou plus c'est une inclusion stricte. Par exemple, le nuplet $t = \langle X_1 = \frac{\sqrt{3}}{2} \Delta^{\max}, X_2 = -\frac{\sqrt{3}}{2} \Delta^{\max}, X_3 = 0 \rangle \in \text{DEVIATION}(\mathcal{X}, m = 0, [0, \sqrt{3} \cdot \Delta^{\max}])$ mais $t \notin \text{SPREAD}(\mathcal{X}, m = 0, [0, (\Delta^{\max})^2])$. En effet, $\left(\frac{\sqrt{3}}{2} \Delta^{\max}\right)^2 + \left(-\frac{\sqrt{3}}{2} \Delta^{\max}\right)^2 + 0^2 = \frac{3}{2} (\Delta^{\max})^2 > (\Delta^{\max})^2$.

8 Résultats expérimentaux

L'objectif de cette section expérimentale est de comparer le filtrage obtenu par les propagateurs décrits aux Sections 4 et 5, avec une implémentation par décomposition suggérée à la Section 3.

20,000 ensembles $\mathcal{X} = \{X_1, \dots, X_{50}\}$ ont été générés. Le domaine d'une variable X est composé des entiers entre le minimum et le maximum d'une paire de nombres générés aléatoirement entre -50 et 50. La contrainte moyenne pour chaque ensemble est $m = 0.5$. La somme des écarts maximum D^{\max} varie entre 200 et 1000. Cet intervalle a été trouvé expérimentalement de sorte qu'avec $D^{\max} = 200$ (resp. 1000) tous les ensembles sont inconsistants (resp. consistants).

La Figure 4 montre les statistiques au noeud racine (aucune recherche n'a été employée). Le nombre d'ensemble inconsistants détectés par les deux approches sont données sur la Figure 4 à gauche. Remarquons que si un exemple est détecté comme inconsistant par décomposition, DEVIATION le détecte aussi. Le pourcentage de filtrage moyen (le nombre de valeurs filtrées divisé par le nombre de valeurs dans les domaines initiaux) sur les ensembles consistants sont représentés sur la Figure 4 à droite.

Le nombre d'ensembles inconsistants détectés est significativement plus important pour les propagateurs présentés que pour une approche par décomposition.

Par exemple, pour $D^{\max} = 500$, DEVIATION détecte 9,619 inconsistances contre 3,628 pour la décomposition. Sur les 10,381 ensembles inconsistants, le pourcentage de filtrage obtenu pour DEVIATION est de 11.8% contre 0.9% pour la décomposition.

Comme expliqué à la Section 6, toutes les inconsistances peuvent ne pas être détectées par nos propagateurs si la moyenne n'est pas entière. C'est le cas dans nos expérience puisque $m = 0.5$. La Figure 5 montre un graphique du pourcentage d'inconsistances détectées par décomposition et par DEVIATION. Presque toutes les inconsistances sont détectées par DEVIATION. Le plus faible pourcentage est de 99.66% pour $D^{\max} = 400$.

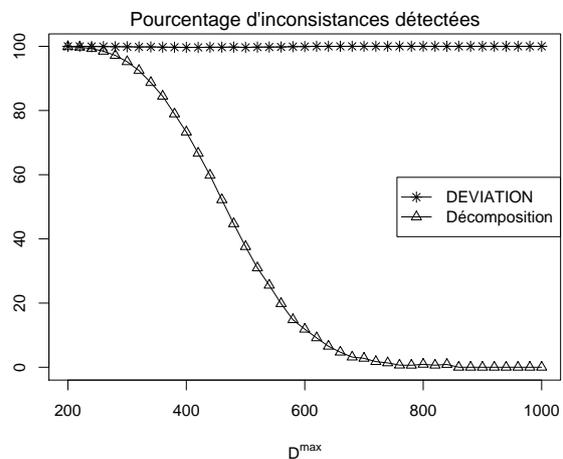


FIG. 5 – Pourcentage d'inconsistance détectées.

La Section 7 introduit une approximation de SPREAD avec DEVIATION. La partie gauche de la Figure 6 montre le nombre d'inconsistances détectées par SPREAD (telle que implémentée dans [7]) comparé au nombre d'inconsistances trouvées par l'approximation. De nombreuses inconsistances restent non détectées mais, comme montré sur la droite de la Figure 6, la propagation utilisant DEVIATION est deux ordres de grandeur plus rapide qu'avec SPREAD sur ces 20,000 instances aléatoires.

9 Conclusion

Ce travail présente DEVIATION, une nouvelle contrainte globale pour équilibrer un ensemble de variables. Cette contrainte est assez similaire à SPREAD [5, 7]. Alors que SPREAD contraint la norme L_2 , DEVIATION contraint la norme L_1 .

Les algorithmes de filtrage que nous introduisons ont une complexité temporelle linéaire par rapport au

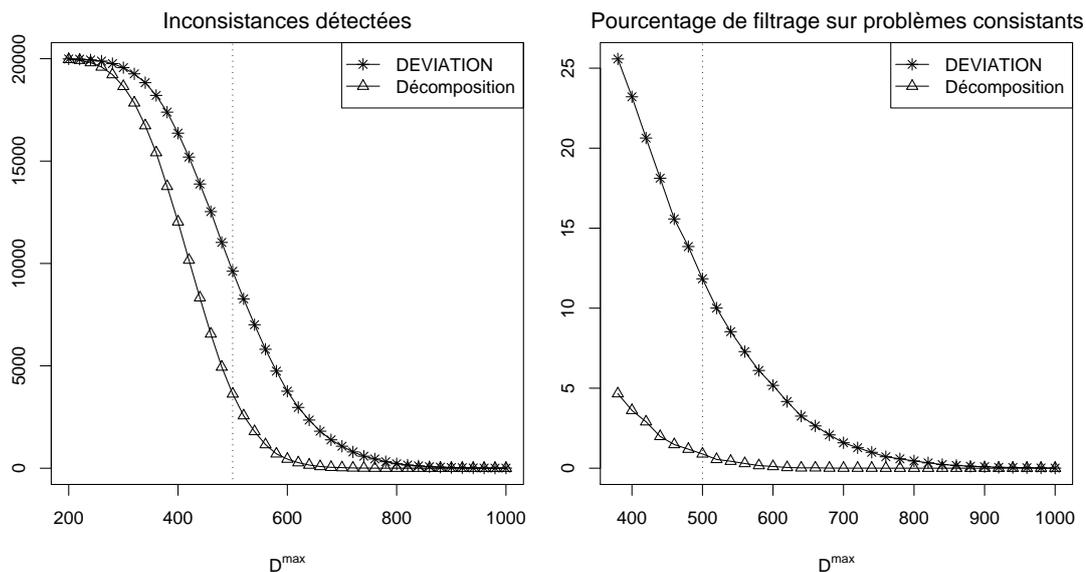


FIG. 4 – Résultats expérimentaux : comparaison de DEVIATION et de la décomposition proposée à la Section 3.

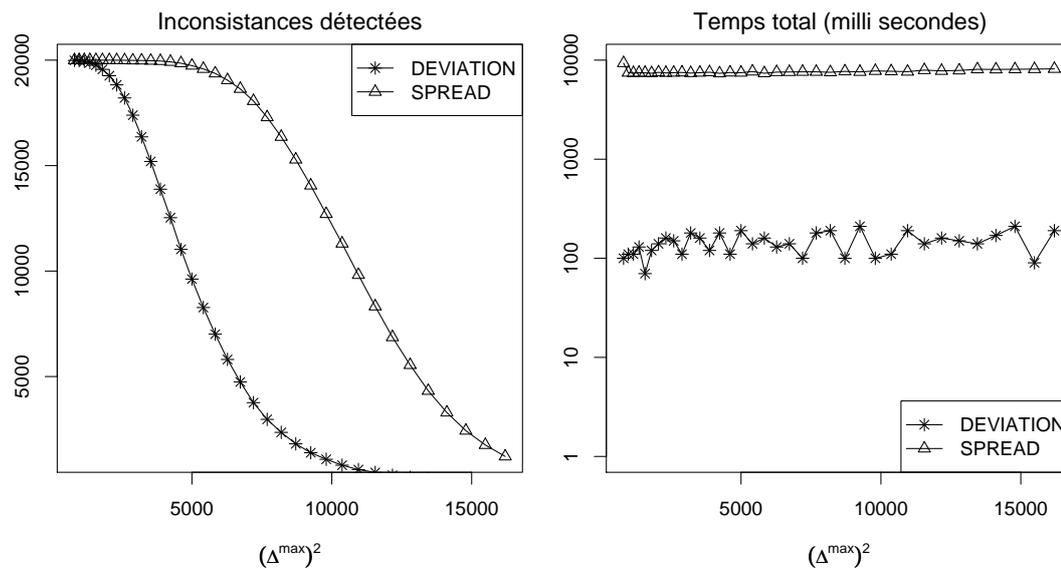


FIG. 6 – Résultats expérimentaux : SPREAD v.s. approximation par DEVIATION.

nombre de variables. Des expériences évaluent l'efficacité de nos propagateurs en termes de filtrage. Une relaxation de SPREAD avec DEVIATION est également introduite. Cette relaxation filtre moins de valeurs que les propagateurs dédiés mais permet néanmoins de réduire le temps de calcul.

Remerciement

Les auteurs aimeraient remercier Julien Hendrickx et Raphaël Jungers pour la preuve du Théorème 4 ainsi que Jean-Noël Monette et Grégoire Dooms pour les nombreuses discussions intéressantes.

Cette recherche a été partiellement supportée par la Région Wallonne (projet Transmaze 516207).

Références

- [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. Bessière, 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. Simplification and extension of the spread constraint. *Third International Workshop on Constraint Propagation And Implementation*, 2006.
- [8] P. Schaus, Y. Deville, P. Dupont, and J.C. Régin. The deviation constraint. *Proceedings of CP-AI-OR*, 4510 :269–284, 2007.
- [9] W.J. van Hoes, 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.