Ensemble et polytope de dépendance, système générateur et ordonnancement Version du 25 février 2004

Rapport interne E/258/CRI

Corinne Ancourt
Isabelle Hurbain
François Irigoin
Nicolas Museux

18 juin 2004

Résumé

Le système d'aide au placement APOTRES de THALES utilise les sommets du polytope de dépendance pour exprimer les contraintes existant entre la date d'exécution de l'itération produisant un élément de tableau et la date d'exécution d'une itération utilisant cet élément de tableau.

L'implémentation actuelle s'appuie sur un théorème de décomposition du polyèdre de dépendance en un produit cartésien de projection sur des plans dont les hypothèses ne sont pas explicitées. La reconstruction du système générateur du polyèdre ou du polytope à partir des systèmes générateurs des projections polygonales n'a pas non plus été explicitée dans les documents antérieurs.

Après avoir revisité le problème de l'ordonnancement avec des exemples très simples, nous étudions pas-à-pas le passage de l'ensemble des dépendances entre itérations au polytope de dépendance, le calcul de ses sommets rationnels, la notion de sommet redondant qui est liée à d'autres modèles, la correction et l'extension de la proposition 9.1.4 qui est à la base d'APOTRES, l'impact du partitionnement paramétrique sur le calcul des dépendances, et enfin le problème du calcul ou de l'approximation des sommets entiers. L'objectif est de ne pas considérer une seule étape comme triviale et donc de définir et de prouver tous les théorèmes sous-jacents.

Nous proposons finalement un calcul direct des coordonnées des sommets entiers d'une sur-approximation du polyèdre de dépendance entier, ce calcul n'étant qu'une extension très simple de la solution définie en 1996 puis raffinée par Nicolas Museux dans sa thèse.

Il reste à l'étendre pour prendre en compte les dimensions infinies et les opérateurs modulo, mais ceci n'est pas nécessaire dans le cadre de DREAM-UP.

Il reste aussi toujours à trouver un moyen pour passer sans approximation ou avec une approximation plus fine d'une formulation rationnelle à une formulation entière.

Ce document est rédigé dans le cadre du contrat RNTL DREAM-UP. Sa lecture suppose une bonne connaissance du contexte, du rapport DRET de 1996 et/ou de la thèse de Nicolas Museux.

1 Introduction

L'analyse des dépendances entre itérations définissant un élément de tableau et itérations utilisant cet élément est une technique bien connue en parallélisation automatique et en ordonnancement. L'approximation de l'ensemble des dépendances par un polyèdre le contenant permet de dériver des contraintes sur les ordonnancements affines valides pour ces dépendances à partir du système générateur du polyèdre.

Cette technique a été utilisée par les systèmes de placement et d'aide au placement d'application temps-réel développés par Christophe Guettier et Nicolas Museux chez THALES en utilisant la programmation par contraintes.

Les contraintes doivent faire apparaitre les variables des différents modèles utilisés. Les contraintes de dépendance, par exemple, font apparaître les variables de partitionnement en sus des fonctions d'accès aux éléments de tableaux qui sont des constantes. Il faut donc trouver un moyen de calculer symboliquement les sommets du graphe de dépendance après partitionnement.

L'implantation actuelle APOTRES produit des ordonnancements non-valides (cf. section 2) dont l'analyse a conduit à un ré-examen des théorèmes justifiant la construction des contraintes d'ordonnancement liées aux dépendances. Les hypothèses d'un théorème clé, la proposition 9.1.4 du rapport DRET[], ne sont pas explicitées et ne sont pas vérifiées dans certains cas.

Le but initial de ce rapport était de voir comment on pourrait corriger cette proposition 9.1.4 par un ou plusieurs théorèmes similaires permettant d'effectuer encore une recomposition exacte de l'ensemble des dépendances ou de son approximation polyédrique à partir d'un ensemble de projections bi-dimensionnelles, ainsi qu'une recomposition du système générateur de l'approximation polyédrique à partir des systèmes générateurs des projections, ou bien comment on pourrait obtenir autrement ce système générateur de l'ensemble des dépendances.

Au vu des exemples et des problèmes rencontrés tant en exactitude qu'en vitesse, il nous a semblé préférable de reprendre le problème à la base comme en 1996.

Les plupart des théorèmes nécessaires sont démontrés sur des espaces affines réels ou rationnels, sans prendre en compte initialement les difficultés supplémentaires dues au fait que l'ensemble des dépendances est construit sur les entiers naturels.

Par ailleurs, nous traitons d'abord le cas où l'ensemble des dépendances est fini et où le polytope correspondant est donc défini par des sommets générateurs. Nous ne prenons pas en compte les opérateurs modulo dans l'adressage des tableaux.

Ce rapport commence par un contre-exemple montrant l'insuffisance de la construction actuelle des contraintes d'ordonnancement (section 2). Il continue avec des exemples suffisamment petits pour pouvoir être traités manuellement et pour valider les résultats d'APOTRES (section 3). Ces exemples pourrait constituer un embryon de suite de tests de non-regression. La section suivante, la section 4, montre comment on passe de l'ensemble des dépendances au polytope des dépendances. Elle contient une analyse de la structure des contraintes du polytope de dépendance, ainsi qu'un algorithme de calcul de ses sommets rationnels. Dans la section 5, on montre que d'autres modèles du placement induisent mathématiquement plusieurs ordres partiels sur les sommets, rationnels ou entiers. Ces ordres fondent la notion expérimentale de sommet redondant, qui n'est pas utilisable en théorie parce qu'elle brise l'indépendance des modèles, mais qui est utile pour accélérer les calculs manuels d'ordonnancement. Nous prouvons ensuite que la proposition 9.1.4 est fausse en général, nous donnons des hypothèses supplémentaires sous lesquelles son principe est préservé, nous la redéfinissons pour ne pas en restreindre l'usage par des hypothèses trop strictes, mais nous n'arrivons pas à obtenir une nouvelle procédure de construction des sommets rationnels ou entiers par cette approche (section 6). Le partitionnement paramétrique est introduit section 7. Si la contrainte architecturale dite SIMD n'est pas utilisée, il faudrait revoir les bornes supérieures utilisées. Le partitionnement accroit les distances pouvant exister entre sommets rationnels et sommets entiers. Ce problème est finalement abordé en section 8. Les problèmes non-traités sont énumérés dans la dernière section, section 9.

2 Contre-Exemple

Ce contre-exemple montre que l'implémentation actuelle conduit à une erreur d'ordonnancement.

Source Nicolas Museux et/ou le compte-rendu de réunion du 22 janvier 2004.

Le code initial:

```
DO i1 = 0, 7
  DO j1 = 0, 7
    B[i1,i2] = ...

DO i2 = 0, 7
  DO j2 = 0, 7
    Y[i2, j2] =f(...,B[*,*],...)
```

Après partitionnement:

```
DO c1i = 0, 3
  DO c1j = 0, 0
    B[4*c1i:4*c1i+3,8*c1j:8*c1j+7] = ...

DO c2i = 0, 0
  DO c2j = 0, 0
    Y[8*c2i: 8*c2i+7, 8*c2j:8*c2j+7] =f(...,B[0:7,0:7],...)
```

L'ensemble des dépendances est inclu dans un espace de dimension 4. Il est donc difficile à représenter a priori, mais le partitionnement de la deuxième tâche, `T_IDCT`, réduit l'ensemble des itérations à un point, $(0, 0)$. Le partitionnement de la première tâche réduit son espace d'itération de deux à une dimension.

L'espace des dépendances est donc mono-dimensionnel. L'ensemble des dépendances est défini par un segment lié à la boucle externe de `T_IN_B`, l'intervalle $[0, 3]$.

Les deux sommets dans l'espace quadri-dimensionnel sont $(0, 0, 0, 0)$ et $(0, 0, 3, 0)$ dans le repère de l'espace des dépendances $(c2i, c2j, c1i, c1j)$.

CA NE CORRESPOND PAS AUX CONTRAINTES DONNEES DANS LE COMPTE_RENDU. D'OU VIENT LE SOMMET DE 8 POUR C2J?

3 Exemples

Une série d'exemples est étudiée manuellement pour vérifier que le comportement d'APOTRES est conforme au comportement prédit par la théorie. Ces exemples pourront être conservés dans une base de tests de non-régression.

Le cas des opérateurs modulo n'est pas pris en compte dans APOTRES. Il n'est pas pris en compte ici. Les dimensions infinies ne sont pas prises en compte non plus.

Aucun exemple ne viole les hypothèses de définition unique des éléments d'un tableau et de définition des éléments utilisés. La réponse d'APOTRES à ces erreurs de spécification devrait cependant être étudiée.

3.1 Cas mono-dimensionnels

Quatre cas mono-dimensionnels de difficultés croissantes ont été étudiés. Dans aucun de ces cas, le consommateur considéré n'utilise tous les éléments produits. Dans aucun de ces cas, le consommateur considéré n'utilise des éléments non produits.

Dans tous les cas, APOTRES a posé les contraintes d'ordonnancement attendues en fonction des dépendances.

3.1.1 Un élement, un producteur et un consommateur

Comme on peut le voir sur la figure 1, il existe une injection des itérations de lecture vers les itérations d'écriture.

```
DO i1 = 0, 39
  X[i1] = ...

DO i2 = 0, 11
  ... = X[i2+10]
```

Les tableaux sont supposés accédés par éléments croissants avec les itérations. On ne peut pas avoir:

```
DO i1 = 0, 39
  X[i1] = ...

DO i2 = 0, 11
  ... = X[-i2+21]
```

ce qui ne permet pas de traiter à un niveau fin une application combinant les premiers et les derniers éléments d'un tableau comme pour :

```
DO i1 = 0, 39
  X[i1] = ...

DO i2 = 0, 19
  ... = X[i2]+X[-i2+39]
```

3.1.2 Un consommateur de plusieurs éléments et un producteur

```
DO i1 = 0, 159
  X[i1] = ...

DO i2 = 0, 11
  ... = X[i2+10+0:i2+10+9]
```

Chaque itération `i2` utilise 10 éléments de X, ce qui est appelé un *motif.* La variable balayant le motif, m_2 par exemple, est définie sur l'intervalle $[0, 9]$. Un décalage de 10, noté K^R par la suite, est aussi appliqué.

Fig. 1 – *Un producteur et un consommateur d'un élément*

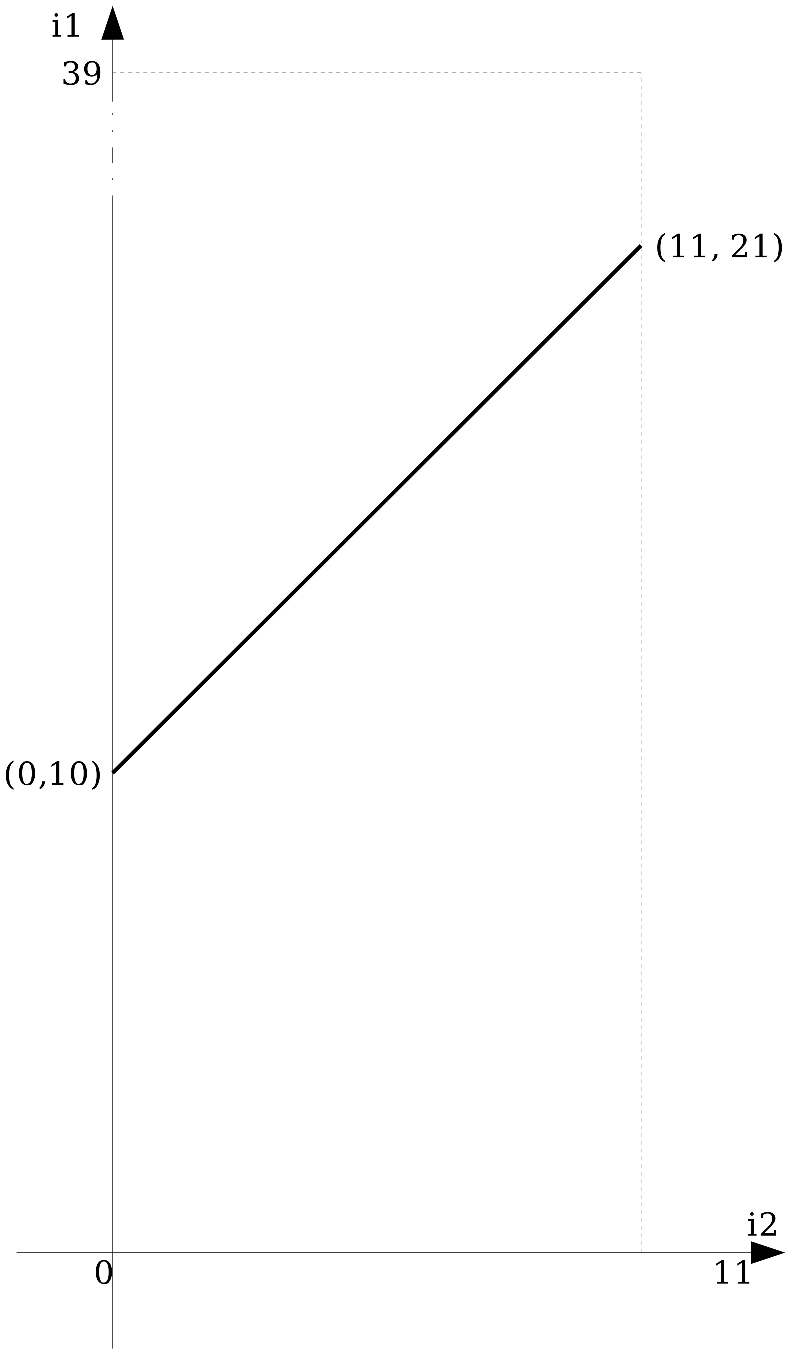
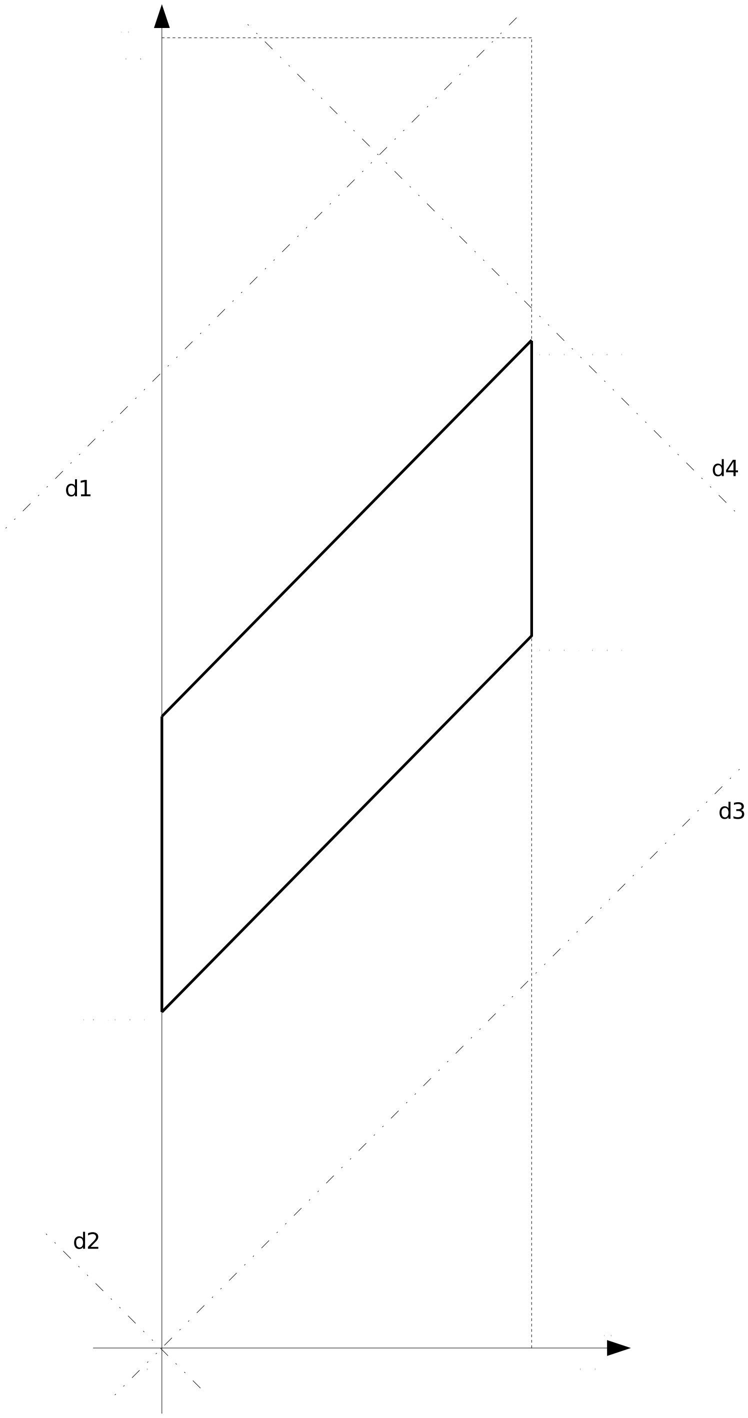


Fig. 2 – *Un producteur et un motif de 10 éléments consommés*



L'explicitation de m_2 permettrait de lui appliquer un coefficient multiplicatif et de sélectionner, par exemple, un élément sur deux ou sur trois.

Quatre ordonnancements valides, d1, d4, d2 et d3 , sont aussi représentés graphiquement sur la figure 2 comme des formes affines donnant le délai entre la production et la consommation d'un élément:

$$
\begin{aligned}
i_1 &\leq i_2 + 29 \\
i_1 &\leq -i_2 + 41 \\
-i_1 &\leq i_2 \\
-i_1 &\leq -i_2
\end{aligned}
$$

Le premier est le plus usuel. C'est le seul qui est valide pour APOTRES parce que les coefficients de i_1 et de i_2 sont tous les deux positifs. Les autres ordonnancements sont donnés à titre d'exemple pour montrer que tous les sommets du polytope de dépendance sont utiles en général, même si ce n'est pas le cas pour APOTRES dont les coefficients d'ordonnancement positifs induisent un ordre sur les sommets du graphe de dépendance et permettent de réduire le nombre de contraintes d'ordonnancement.

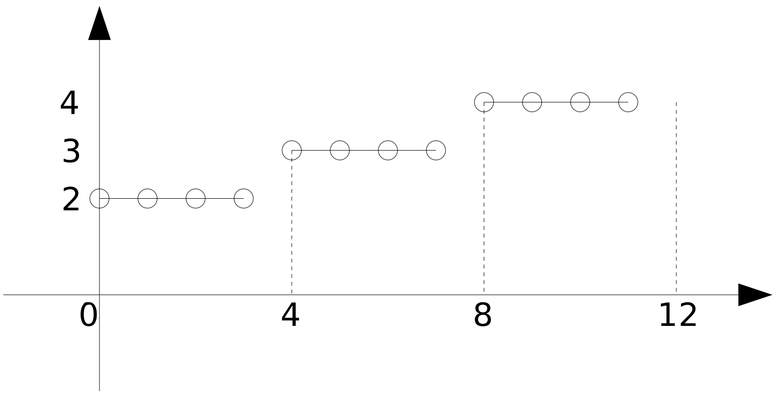
Le second ordonnancement consiste à effectuer les itérations i_2 en ordre inverse, après avoir attendu un délai suffisant. Le troisième consiste à effectuer les itérations de i_1 à des instants négatifs et les itérations de i_2 à des instants positifs. La première tâche est donc intégralement terminée avant que la seconde ne commence.

Le quatrième ordonnancement consiste à effectuer les deux tâches en ordre inverse. Comme les iterations i_1 dont a besoin une itération i_2 sont plus grandes, elles sont exécutées avant.

Chacun de ces quatres ordonnancements appartient à un ensemble d'ordonnancements différents. Dans chacun de ces ensembles, de nombreux ordonnancements sont strictement équivalents quant aux ordres partiels ou totaux induits.

3.1.3 Un producteur de plusieurs éléments et un consommateur

FIG. 3 – *Un producteur de plusieurs éléments et un consommateur*



```
REAL X[0:]
DO i1 = 0, 39
  X[4*i1:4*i1+3] = ...

DO i2 = 0, 11
  ... = X[i2+8]
```

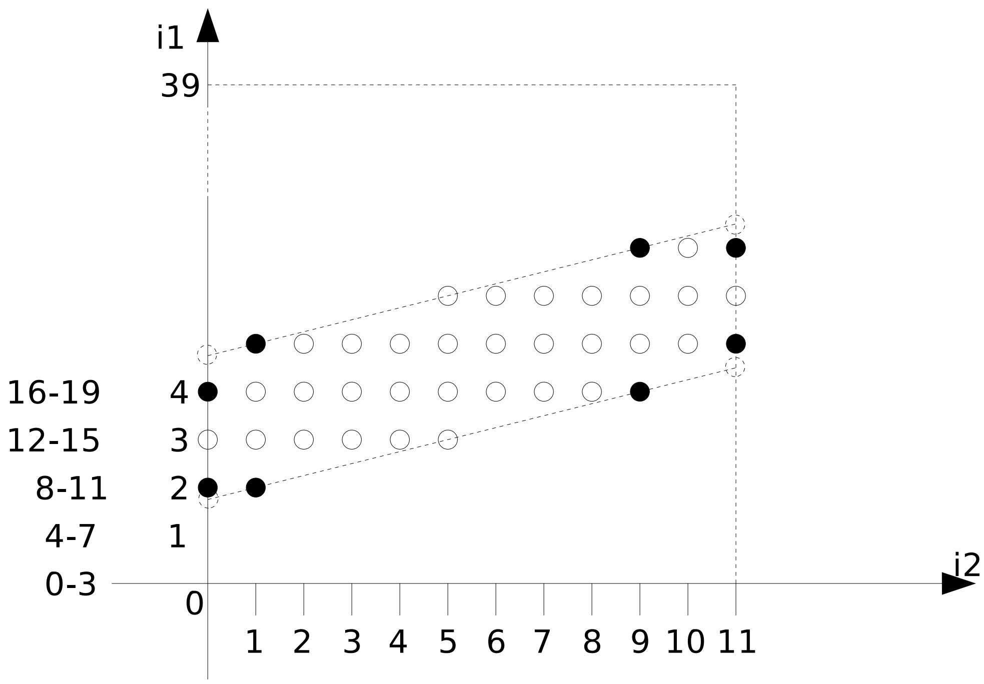
Chaque itération `i1` définit 4 éléments de X, ce qui est appelé un *motif*. La variable balayant le motif, m_1 par exemple, est définie sur l'intervalle $[0, 3]$. Ce motif doit être contigu. Sa borne supérieure, 3, est liée au coefficient multiplicatif de `i1`, 4, afin d'assurer la définition de chaque élément du tableau. Les éléments sont utilisés un par un, ce qui produit un phénomène périodique de marches d'escalier sur la figure 3.

3.1.4 Un producteur et un consommateur d'éléments

Maintenant, les éléments sont produits quatre par quatre, puis utilisés dix par dix avec un décalage de un seulement.

L'ensemble des dépendances du code suivant est représenté sur la figure 4. Les dépendances sont représentées par des cercles vides ou pleins. Les cercles pleins représentent les sommets entiers du polyèdre de dépendance. Les contraintes rationnelles et les sommets rationnels sont aussi représentés.

FIG. 4 – *Un producteur et un consommateur d'éléments*




```
DO i1 = 0, 39
  X[4*i1:4*i1+3] = ...
```

```
DO i2 = 0, 11
  ... = X[i2+10+0:i2+10+9]
```

On note que le nombre de sommets entiers est deux fois plus grands que le nombre de sommets rationnels et que la dérivation des sommets entiers à partir des sommets rationnels n'est pas évidente. Malheureusement, APOTRES ne peut travailler que sur des sommets entiers. Le problème du passage des sommets rationnels aux sommets entiers est étudié en section 8.

3.1.5 Conclusion sur les cas monodimensionnels

APOTRES a correctement calculé les polytopes de dépendances, qui sont dans ce cas des polygones, et posé toutes les contraintes. Ce n'est pas étonnant puisque le polygone de dépendance est égal à son unique projection 2-D.

Il faudrait aussi voir comment APOTRES réagit si la tâche utilisatrice fait appel à des éléments non définis. La compatibilité de la spécification avec les hypothèses sur les applications est-elle vérifiée?

3.2 Cas bi-dimensionnel

3.2.1 Un consommateur de plusieurs éléments et un producteur

```
DO i1 = 0, 159
  DO j1 = 0, 159
    X[i1, j1] = ...

DO i2 = 0, 11
  DO j2 = 0, 11
    ... = X[i2+10+0:i2+10+9, j2+10+0:j2+10+9]
```

Quatre des seize sommets du polytopes de dépendances seulement ont un impact sur les ordonnancements respectant l'ordre lexicographique[1]: $(0, 0, 19, 19)$, $(0, 11, 19, 30)$, $(11, 0, 30, 19)$ et $(11, 11, 30, 30)$, en donnant les coordonnées dans l'ordre `(i2, j2, i1, j1)`.

L'ordonnancement libre séquentiel au plus tôt, compatible avec l'ordre lexicographique, est donné pour la première boucle par:

$$d^W(i_1, j_1) = 160i_1 + j_1$$

L'ordonnancement au plus tôt de la deuxième boucle:

$$d^R(i_2, j_2) = 160i_2 + j_2 + 3059$$

s'en dérive en utilisant les contraintes imposées par les quatre sommets utiles et en prenant en compte la numérotation des tâches qui permet de remplacer une inégalité stricte par une inégalité simple:

$$160 \times 19 + 19 \leq \gamma^R \Longrightarrow 3059 \leq \gamma^R$$

1. Cette sélection est expliquée section 5. L'ordre lexicographique est une condition suffisante, mais le fait que les coefficients des ordonnancements soient des constantes positives et donc les ordonnancements des fonctions croissantes induit aussi un ordre partiel sur les sommets du polytope de dépendances.

$$160 \times 19 + 30 \leq 11\delta_2^R + \gamma^R \Longrightarrow 1 \leq \delta_2^R$$

$$160 \times 30 + 19 \leq 11\delta_1^R + \gamma^R \Longrightarrow 160 \leq \delta_1^R$$

$$160 \times 30 + 30 \leq 11\delta_1^R + 11\delta_2^R + \gamma^R \Longrightarrow 1771 \leq 11\delta_1^R + 11\delta_2^R$$

Comme $\delta_1^R \geq 12\delta_2^R$ et $\delta_2^R > 0$:

$$1771 \leq 12\delta_1^R - \delta_2^R \leq 12\delta_1^R \Longrightarrow 147 \leq \delta_1^R$$

Cet exemple a été soumis à APOTRES. Il a permis de vérifier que APOTRES ne posait pas les contraintes relatives aux sommets 2 et 3 et qu'il passait beaucoup trop de temps dans sa recherche de solutions pour un exemple trivial, même si le partitionnement est imposé.

Une fois imposé les coefficients δ^R et δ^W de l'ordonnancement, il a fallu laisser γ^R libre parce que APOTRES déduit d'autres contraintes, non encore identifiées, qu'il augmenter le délai entre W et R (contribution NICOLAS MUSEUX).

Question: est-ce que l'ordre lexicographique sur W et l'hypothèse des accès croissants implique l'ordre lexicographique sur R? Dans le cas général, non, parce que W peut produire tout le tableau X en une unique itération via les motifs et que l'ordre sur R est alors libre. D'une manière plus générale, pour des boucles toutes finies, il suffit de choisir $\gamma^R \geq \delta^W I^W$ pour supprimer les contraintes sur δ^R.

3.3 Conclusion sur les exemples

Les exemples sont indispensables pour construire une suite de non-régression.

La PPC permet d'imposer la ou les solutions trouvées manuellement et de les vérifier ou de les infirmer rapidement. En cas de contradiction, il est plus difficile de trouver l'origine du problème.

4 Passage de l'ensemble des dépendances au polytope de dépendance

On considère d'abord le cas où l'ensemble des dépendances est fini et où il peut être approximé sans perte d'information par un polyèdre obtenu par une fermeture convexe parce que les ordonnancements considérés sont affines.

Le modèle de partitionnement n'est pas pris en compte.

4.1 Hypothèse et notations

On considère les dépendances entre les itérations w d'un nid W de n_W boucles produisant tous[2] les éléments d'un tableau X et les itérations r d'un nid R de n_R boucles utilisant certains des éléments de ce tableau.

Le tableau X de dimension n_X est défini hypercube par hypercube, ces hypercubes étant tous identiques et parallèles aux axes[3]. Les hypercubes peuvent être réduits à un élément ou à

2. Ceci n'est pas vérifié par APOTRES (?).
3. De tels hypercubes sont appelés *motifs*.

n'importe quel hypercube de dimension intermédiaire. La famille d'hypercubes représente une partition du tableau. Chaque hypercube est appelé un *motif*.

Le tableau X de dimension n_X est utilisé image d'hypercube par image d'hypercube, ces hypercubes étant tous identiques. Les fonctions "image" sont des matrices diagonales entières. Les hypercubes peuvent être réduits à un élément ou à n'importe quel hypercube de dimension intermédiaire en réduisant l'épaisseur d'une ou plusieurs dimensions à un élément. Mais la famille d'hypercubes ne représente pas forcément une partition du tableau. Les hypercubes peuvent avoir des intersections non-vides. De plus, leur union peut ne pas couvrir l'ensemble du tableau. On suppose néanmoins qu'elle est inclue dans cet ensemble[4], sans même avoir à appliquer un opérateur modulo.

Le tableau X est défini par une fonction T_W associée au nid de boucles W et utilisé par une fonction T_R associée au nid de boucles R. L'ensemble des itérations w est noté I^W et l'ensemble des itérations r est noté I^R. Les éléments du tableau X sont repérés par des vecteurs x ou ϕ. Ils dépendent des itérations w ou r ainsi que des coordoonnées relatives des éléments des hypercubes accédés, m et n. Les vecteurs m et n prennent leurs valeurs dans M^W et M^R, des hypercubes de dimension n_X aussi définis par $[0..M^W-1]$ et $[0..M^R-1]$ en surchargeant M^W et M^R pour simplifier les écritures ultérieures.

En définition de tableau, l'élément considéré est:

$$\phi = \Omega^W_{pav} w + \Omega^W_{fit} m + K^W = \Omega^W_{pav} w + m$$

parce que Ω^W_{fit} est la matrice identité et que K^W est nul, la borne inférieure des dimensions de tableau étant nulle. On a de plus:

$$\forall k \in [1, n_X] \exists i \in [1, n_W] ({\Omega_{pav}}^W_{k,i} = M^W_k \vee {\Omega_{pav}}^W_{k,i} = 0) \wedge (\forall i' \in [1, n_W] i' \neq i \Longrightarrow {\Omega_{pav}}^W_{k,i} = 0)$$

pour qu'un hypercube de ϕ soit bien défini. Cette formule devrait être clarifiée si possible. Chaque dimension du tableau peut soit être accédée par l'itérateur et une composante du motif, éventuellement limitée à un élément, soit par le motif seul.

En utilisation de tableau, l'accès est défini par:

$$\phi = \Omega^R_{pav} r + \Omega^R_{fit} n + K^R$$

où Ω^R_{pav}, Ω^R_{fit} et K^R sont des matrices ou vecteurs à coefficients entiers positifs ou nuls.

Le nid de boucle W ne définit qu'un unique tableau X et le tableau X n'est défini que par ce nid de boucle. Le nid de boucle R utilise plusieurs tableaux, voire plusieurs fois le même tableau via des références différentes. On ne s'intéresse ici qu'à la référence de définition et à une des références d'utilisation. Le numéro de référence n'est pas explicité dans ce document.

Les matrices Ω_{fit} sont des matrices diagonales. Les motifs ont une dimension inférieure à la dimension du tableau correspondant. On peut en augmenter la dimension en prenant des épaisseurs nulles. L'ordre d'énumération des dimensions des motifs n'a pas d'impact. Il peut être pris égal à l'ordre d'énumération des dimensions du tableau.

4.2 Ensemble des dépendances

Un ensemble de dépendance est défini entre les itérations d'un nid de boucles W définissant un tableau X et un nid de boucle R contenant une ou plusieurs références à ce tableau. Un ensemble de dépendance devrait être dénoté soit par le tableau X soit par la tâche W puisque

4. Ceci n'est pas vérifié par APOTRES.

tableaux et tâches de définition sont en bijection, soit par la tache R seule si une seule occurence de chaque tableau est autorisée, soit par la tâche R et un numéro de référence sinon.

Dans ce rapport, le nom du tableau concerné est obtenu implicitement par la mention de W. Le nom du tableau concerné n'est pas rappelé pour qualifier les matrices d'accès de la tâche R. Le rang de l'accès n'est pas non plus spécifié. Ce contexte implicite permet d'alléger les notations, mais une implémentation devrait l'expliciter.

L'ensemble de dépendance entre les tâches W et R est défini par:

$$\begin{aligned} D^{W,R} \; = \; & \Big\{(r,w) \in \mathbb{N}^{n_R} \times \mathbb{N}^{n_W} | w \in I^W, r \in I^R, \exists m \in M^W, \exists n \in M^R \; t.q. \\ & \Omega^W_{pav} w + m = \Omega^R_{pav} r + \Omega^R_{fit} n + K^R \Big\} \end{aligned} \tag{1}$$

Les matrices d'accès Ω^W_{pav}, Ω^R_{pav} et Ω^R_{fit} sont à coefficients entiers positifs. Les ensembles M^W et M^R sont des produits d'intervalles de bornes inférieures nulles et de bornes supérieures entières strictement positives. Ces bornes supérieures peuvent aussi être notées comme des vecteurs M^W et M^R, à condition d'ouvrir les intervalles à droite: $m \in [0, M^W[$ et $n \in [0, M^R[$.

Les matrices Ω^R_{pav} et Ω^R_{fit} et le vecteur M^R devraient être qualifiés par le rang de l'accès dans le nid de boucles R et/ou par le nom du tableau concerné. La première qualification est plus simple ou plus robuste parce qu'elle permet de référencer une ligne et une colonne d'une matrice sans avoir à référencer toute la matrice.

4.3 Approximation convexe de l'ensemble des dépendances

On montre dans cette section que l'ensemble des dépendances peut être remplacé par le polyèdre des dépendances, sa fermeture convexe sur les rationnels (ou les réels), sans réduire l'ensemble des ordonnancements affines valides.

4.3.1 Fermeture convexe des dépendances et ordonnancements

Une paire de formes affines est un ordonnancement valide si les délais associés à toutes les dépendances par ces formes sont strictement positifs, ou simplement positifs si on prend en compte APOTRES, son système de numérotation des nids de boucles et son impact sur l'ordonnancement.

$$\forall w \forall r D^{W,R}(r,w) \Longrightarrow \delta^W w + \gamma^W \leq \delta^R r + \gamma^R \Longrightarrow \delta^R r + \gamma^R - \delta^W w - \gamma^W \geq 0$$

Théorème 0.4-1 : Si un ordonnancement affine est valide pour deux vecteurs de dépendances $d_1 = (r_1, w_1)$ et $d_2 = (r_2, w_2)$, il est valide pour toutes leurs combinaisons linéaires $\lambda d_1 + (1 - \lambda) d_2$, avec $\lambda \in [0, 1]$.

Preuve: Le délai entre les itérations de définition et d'utilisation est aussi une forme affine définie par:

$$(\delta^{R,W}, \gamma^{R,W}) = ((\delta^R, -\delta^W), \gamma^R - \gamma^W)$$

Comme $\lambda \in [0, 1]$, λ et $1 - \lambda$ sont positifs. On peut donc multiplier les inégalités assurant la validité de d_1 et de d_2 pour obtenir:

$$\lambda \delta^{R,W} d_1 + \lambda \gamma^{R,W} \geq 0$$

et

$$(1 - \lambda) \delta^{R,W} d_2 + (1 - \lambda) \gamma^{R,W} \geq 0$$

qui peuvent être additionnées et factorisées pour donner l'inégalité:

$$\delta^{R,W}(\lambda d_1 + (1-\lambda)d_2) + \gamma^{R,W} \geq 0$$

qui montre que la combinaison est une dépendance respectée par l'ordonnancement.

QED

Définition: un ordonnancement affine o est valide pour un ensemble de dépendance D, si pour tout élément d de D on a $o.d > 0$.

$$valide_D(o) \Longleftrightarrow \forall d \in D \quad o.d > 0$$

Corollaire: L'ensemble O des ordonnancements affines o valides par rapport à l'ensemble des dépendances D est égal à l'ensemble O' des ordonnancements affines o' valides par rapport à la fermeture convexe $\bar{D}$ de l'ensemble des dépendances.

Preuve: Quelque soit o' de O', quelque soit d de D, d appartient à $\bar{D}$ par construction et donc $o'.d > 0$. L'ordonnancement o' est donc valide pour D par définition.

Quelque soit o de O, quelque soit d' de $\bar{D}$, soit d' appartient à D et $o.d' > 0$, soit d' est une combinaison linéaire finie d'élément d_i de D. Comme pour tout d_i, $o.d_i > 0$, on a $o.d' > 0$. L'ordonnancement o est donc valide pour $\bar{D}$.

QED

Cette deuxième preuve est trop rapide. Le respect du terme constant impose bien $\sum_i \lambda_i = 1$ alors que cette condition n'est pas utilisée.

4.3.2 Définition du polyèdre de dépendance

L'ensemble des dépendances entre les tâches W et R est sur-approximé polyédriquement en négligeant les problèmes de divisibilité entière et en éliminant les quantificateurs existentiels par:

$$\begin{aligned} \bar{D}^{W,R} \quad = \quad & \Big\{(w,r) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | w \in I^W, r \in I^R \ t.q. \\ & \Big(\Omega^W_{pav} w \geq \Omega^R_{pav} r + K^R - M^W + 1\Big) \wedge \Big(\Omega^W_{pav} w \leq \Omega^R_{pav} r + \Omega^R_{fit}(M^R - 1) + K^R\Big) \ \Big\} \end{aligned} \quad (2)$$

Chaque égalité de l'ensemble des dépendances est remplacée par une paire d'inéquations, dont les projections sur des plans sont appelées D_{sup} et D_{inf} dans la thèse de Nicolas Museux.

En présence de motifs dégénérés, $M^W = 1$ et $M^R = 1$, les deux inégalités sont équivalentes à une égalité, $\Omega^W_{pav} w = \Omega^R_{pav} r + K^R$, mais il est plus simple de continuer avec une paire d'inégalités pour ne pas créer de cas particuliers.

4.3.3 Preuve de l'élimination des quantificateurs de motifs

L'élimination des quantificateurs existenciels, implicite dans la section précédente, peut se faire de manière formelle en considérant les deux ensembles des éléments de tableaux ϕ lus et écrits, en montrant qu'ils sont tous les deux définis par ou inclus dans des intervalles parce que M^W, M^R et Ω^R_{fit} sont positifs, en écrivant les quatre contraintes correspondantes sur ϕ, puis en projetant ϕ par Fourier-Motzkin pour associer la borne inférieure de chaque ensemble à la borne supérieure de l'autre.

Toute cette section n'est que la preuve de la section précédente.

Soit $X^R(r)$ une fonction associant à chaque itération r l'ensemble des éléments de X qui sont lus:

$$X^R(r) = \left\{\phi | \exists n \in M^R, \quad r \in I^R \quad \phi = \Omega^R_{pav} r + \Omega^R_{fit} n + K^R\right\}$$

et $X^W(w)$ une fonction associant à chaque itération w l'ensemble des éléments qui sont écrits:

$$X^W(w) = \left\{\phi | \exists m \in M^W, \quad w \in I^W \quad \phi = \Omega^W_{pav} w + m\right\}$$

Les images de ces fonctions sont vides quand les itérations arguments n'appartiennent pas à l'ensemble d'itérations. On suppose que les spécifications de l'application sont correctes et donc que tous les ϕ appartiennent aux éléments déclarés du tableau X. L'équation 1 peut être réécrite sous la forme:

$$D^{W,R} = \left\{(w,r) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | X^R(r) \cap X^W(w) \neq \emptyset\right\}$$

qui est équivalente à:

$$D^{W,R} = \left\{(w,r) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | \exists \phi \in X^R(r) \phi \in W^W(w)\right\}$$

Par hypothèse, la tâche W remplit continuement le tableau X et $X^W(w)$ a pour image un intervalle, éventuellement vide:

$$\begin{aligned} X^W(w) &= \left\{\phi | \exists m \in M^W, \quad w \in I^W \quad \phi = \Omega^W_{pav} w + m\right\} \\ &= \left\{\phi | w \in I^W \wedge \phi \in [\Omega^W_{pav} w, \Omega^W_{pav} w + M^W - 1]\right\} \end{aligned}$$

La tâche R ne lit pas le tableau X continument mais elle le lit de manière monotone strictement croissante parce qu'on a fait les hypothèses $M^R > 0$ et $\Omega^R_{fit} \geq I$.

$$\begin{aligned} X^R(r) &= \left\{\phi | \exists n \in M^R, \quad r \in I^R \quad \wedge \quad \phi = \Omega^R_{pav} r + \Omega^R_{fit} n + K^R\right\} \\ &\subseteq \bar{X}^R(r) = \left\{\phi | r \in I^R \quad \wedge \quad \phi \in [\Omega^R_{pav} r + K^R, \Omega^R_{pav} r + \Omega^R_{fit}(M^R - 1) + K^R]\right\} \end{aligned} \quad (3)$$

Les extrema entiers des deux intervalles $X^W(w)$ et $\bar{X}^R(r)$ sont atteints par définition, du moins quand ces intervalles ne sont pas vides.

Théorème 0.4-2 : si l'intersection de $X^W(w)$ et $\bar{X}^R(r)$ est non vide, l'intersection de $X^W(w)$ et $X^R(r)$ est non vide ou bien il existe deux itérations w_1 et w_2 telles que $w_1 < w < w_2$ et les intersections de $X^W(w_1)$ et $X^R(r)$ d'une part, et de$X^W(w_2)$ et $X^R(r)$ d'autre part sont non vides.

Preuve: Soit ϕ élément de l'intersection non-vide de $X^W(w)$ et $\bar{X}^R(r)$. Si ϕ appartient à l'intersection de $X^W(w)$ et $X^R(r)$ le théorème est prouvé. Sinon, par définition,

$$\Omega^R_{pav} r + K^R < \phi < \Omega^R_{pav} r + \Omega^r_{fit}(M^R - 1) + K^R$$

Les deux points extrémaux, $\Omega^R_{pav} r + K^R$ et $\Omega^R_{pav} r + \Omega^r_{fit}(M^R - 1) + K^R$ sont calculés par deux itérations w_1 et w_2 puisque tous les éléments du tableau X sont définis par W. Comme les éléments de tableaux sont des fonctions croissantes des itérations, $w_1 < w < w_2$ (UN PEU RAPIDE, C'EST PLUTOT L'INVERSE QUI EST VRAI).

N'EST-CE PAS UN DOUBLON?

Théorème 0.4-3 : la sur-approximation de l'ensemble des dépendances $\bar{D}^{W,R}$ a le même ensemble d'ordonnancements valides que l'ensemble des dépendances $D^{W,R}$

$$\bar{D}^{W,R} = \left\{(r,w) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | \exists \phi \in \bar{X}^R(r) \quad r \in I^R \wedge w \in I^W \wedge \phi \in W^W(w)\right\}$$

Preuve: si (r,w) appartient à $\bar{D}^{W,R}$, alors (r,w) appartient à $D^{W,R}$ ou bien il existe (r,w_1) et (r,w_2) dans $D^{W,R}$ et il existe λ tel que $(r,w) = \lambda(r,w_1) + (1-\lambda)(r,w_2)$ parce que $w_1 < w < w_2$.

CETTE PREUVE EST CORRECTE POUR DES ESPACES D'ITERATIONS MONO-DIMENSIONNELS PARCE QU'ON PEUT CALCULER LAMBDA: $\lambda = \frac{w-w_2}{w_1-w_2}$. LE THEOREME PRECEDENT PLACE w DANS UN HYPERCUBE D'ITERATIONS EXTREMES. IL FAUDRAIT PROUVER QUE TOUS LES SOMMETS DE CET HYPERCUBE CORRESPONDENT A DE VRAIES DEPENDANCES. w SERAIT ALORS EFFECTIVEMENT UNE COMBINAISON DE CES SOMMETS.

La sur-approximation polyédrique des dépendances $\bar{D}^{W,R}$ peut être réécrite:

$$\begin{aligned}
\bar{D}^{W,R} \quad = \quad & \{(r,w) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | \exists \phi \in \mathbb{N}^{n_X} \\
& \Omega^R_{pav} r + K^R \leq \phi \leq \Omega^R_{pav} r + \Omega^r_{fit}(M^R - 1) + K^R \\
& \wedge \quad \Omega^W_{pav} w \leq \phi \leq \Omega^W_{pav} w + M^W - 1 \\
& \wedge \quad 0 \leq r \leq I^R - 1 \\
& \wedge \quad 0 \leq w \leq I^W - 1 \ \}
\end{aligned} \tag{4}$$

La variable ϕ peut être projetée par Fourier-Motzkin sans introduire de nouveaux points entiers dans $\bar{D}^{W,R}$ parce que tous ses coefficients sont égaux à un[] (Ancourt/Pugh).

$$\begin{aligned}
\bar{D}^{W,R} \quad = \quad & \{(r,w) \in \mathbb{N}^{n_W} \times \mathbb{N}^{n_R} | \\
& \Omega^R_{pav} r + K^R \leq \Omega^W_{pav} w + M^W - 1 \\
& \wedge \quad \Omega^W_{pav} w \leq \Omega^R_{pav} r + \Omega^r_{fit}(M^R - 1) + K^R \\
& \wedge \quad 0 \leq r \leq I^R - 1 \\
& \wedge \quad 0 \leq w \leq I^W - 1 \ \}
\end{aligned} \tag{5}$$

En conclusion, la définition 2 est une condition nécessaire non suffisante de dépendance. Il s'agit donc d'une sur-approximation des dépendances, appliquant une fermeture convexe sur les itérations r. Comme on l'a vu (Théorème 1), une telle fermeture n'a pas d'impact sur les ordonnancements affines valides.

4.4 Étude des contraintes du polyèdre de dépendances

On suppose que les matrices d'accès Ω^W_{pav} et Ω^R_{pav} ont au plus un élément non nul par ligne, mais qu'elles peuvent en voir un nombre quelconque par colonne ce qui n'est pas l'hypothèse APOTRES habituelle.

D'après la définition 5, chaque dimension k du tableau X peut créer de nouvelles contraintes sur les itérations par rapport à leurs domaines de définitions I^W et I^R. Pour chaque ligne k, quatre cas disjoints sont possibles en fonction des coefficients de Ω^R_{pav} et de Ω^W_{pav}.

- soit une paire de couplages entre une coordonnée j de w et une coordonnée i de r si les lignes k des deux matrices ont chacune un coefficient non nul; à cause des hypothèses faites sur les accès, au plus un coefficient est non nul.

$$
\begin{aligned}
{\Omega_{pav}}^R_{k,i} r_i + K^R_k &\le {\Omega_{pav}}^W_{k,j} w_j + M^W_k - 1 \\
{\Omega_{pav}}^W_{k,j} w_j &\le {\Omega_{pav}}^R_{k,i} r_i + {\Omega_{fit}}^R_{k,k}(M^R_k - 1) + K^R_k \\
0 \le r_i &\le I^R_i - 1 \\
0 \le w_j &\le I^W_j - 1 \qquad (6)
\end{aligned}
$$

- soit une paire de nouvelles contraintes monodimensionnelles sur une coordonnée j de w due au motif de lecture M^R si la ligne k de la matrice Ω^W_{pav} contient un élément j non nul ${\Omega_{pav}}^W_{k,j}$; ces deux contraintes sont redondantes si la tâche R utilise tous les éléments définis de X

$$
{\Omega_{pav}}^W_{k,j} w_j \ge K^R_k - M^W_k + 1 \ \wedge \ {\Omega_{pav}}^W_{k,j} w_j \le {\Omega_{fit}}^R_{k,k}(M^R_k - 1) + K^R_k \qquad (7)
$$

- soit une paire de contraintes monodimensionnelles sur une coordonnée i de r due au motif de lecture M^W si la ligne k de la matrice Ω^R_{pav} contient un élément non nul; cette contrainte doit être redondante parce que tout élément lu est supposé défini ($\forall r \in I^R, \forall m^R \in M^R, M^W \ge \Omega^R_{pav} r + \Omega^R_{fit} m^R + K^R \ge K^R \ge 0$), l'inverse n'étant pas toujours vrai;

$$
M^W_k - 1 \ge {\Omega_{pav}}^R_{k,i} r_i + K^R_k \ \wedge \ 0 \le {\Omega_{pav}}^R_{k,i} r_i + {\Omega_{fit}}^R_{k,k}(M^R_k - 1) + K^R_k \qquad (8)
$$

- soit une paire de contraintes constantes trivialement satisfaites si les deux lignes k des matrices d'accès Ω^W_{pav} et Ω^R_{pav} sont nulles, puisque tous les coefficients sont positifs et puisque tous les éléments lus doivent avoir été définis[5] ($M^W \ge \Omega^R_{fit} M^R + K^R$).

$$
M^W_k \ge K^R_k \ \wedge \ 0 \le {\Omega_{fit}}^R_{k,k} M^R_k + K^R_k \qquad (9)
$$

On en conclut que le polyèdre de dépendance est défini par $n_W + n_R$ paires de contraintes monodimensionnelles induites par les ensembles d'itérations ou par une contrainte d'accès monodimensionnelle et par au plus $2 \times n_X$ contraintes bi-dimensionnelles plus strictes que les précédentes puisqu'elles impliquent les ensembles d'itérations. Mais le nombre d'inégalités non-redondantes définissant le polyèdre de dépendance est toujours $2 \times (n_W + n_R)$.

Théorème 0.4-4 : la dimension d'un tableau X est supérieure ou égale au nombre de boucles du nid W le définissant, $n_X \ge n_W$.

Preuve: sous les contraintes dues à la définition unique de chaque élément de X,

$$
\forall j \in [1..n_W] \exists k \in [1..n_X] \ t.q. \ {\Omega_{pav}}^W_{k,j} \ne 0
$$

à chaque colonne j est associée une ligne k.

Deux colonnes ne peuvent pas être associées à la même ligne parce que dans chaque ligne k de Ω^W_{pav} une colonne au plus est non nulle. La fonction des lignes vers les colonnes est injective. Le nombre de lignes doit donc être plus grand ou égal au nombre de colonne, $n_X \ge n_W$.

QED

Théorème 0.4-5 : Sous l'hypothèse de bonne définition de tous les éléments utilisés et sous l'hypothèse de structure de Ω^W_{pav}, le nombre minimal d'inégalités non-redondantes définissant le polyèdre de dépendance $D^{W,R}$ est toujours inférieur à $2 \times (n_W + n_R)$.

5. APOTRES ne pose peut-être pas la contrainte demandant que tous les éléments lus sont définis (NICOLAS MUSEUX).

Preuve: le nombre d'inégalités présentes dans la définition est $2 \times (n_X + n_W + n_R)$.

On conserve les $2n_R$ contraintes dues à I^R.

Comme chaque élément du tableau défini par la tâche W n'est défini qu'une fois, les n_W coordonnées de w doivent apparaître deux fois chacune dans les $2n_X$ contraintes d'accès. Si une des contraintes dues à I^W n'était pas redondante, un élément de X serait utilisé mais non défini. Les $2n_W$ contraintes dues à I^W sont donc redondantes, mais on garde $2n_W$ contraintes dues aux accès à X.

S'il reste d'autres contraintes dues aux accès à X ($n_X > n_W$), elles ne peuvent pas faire intervenir w. Les contraintes qui ne font pas intervenir r non plus sont constantes et donc redondantes. Les contraintes qui font intervenir r ne peuvent pas restreindre l'espace d'itération I^R (sémantique des bornes de boucles). Elles sont donc redondantes par rapport aux contraintes dues à I^R.

Le nombre maximal de contraintes non-redondantes est donc $2 \times (n_R + n_W)$.

QED

Quand les motifs sont réduits à un élément dans certaines dimensions, une paire d'inégalités peut être équivalente à une égalité.

4.5 Coordonnées des sommets rationnels du polytope de dépendance

Les sommets du polytope de dépendance appartiennent à un espace de dimension $n_R + n_W$. Traditionnellement dans l'iconographie APOTRES, les n_R composantes correspondant aux itérations r de I^R sont les premières composantes, et les n_W composantes correspondant aux itérations w de I^W sont les dernières composantes. Les sommets s du polytope de dépendance peuvent donc s'écrire (r, w) où r est une itération utilisant un des éléments de tableau définis par l'itération w.

Du fait du nombre de contraintes, $2 \times (n_R + n_W)$ au plus (cf. ci-dessus), le nombre de sommets est toujours au plus de $2^{n_R+n_W}$. Dans un espace de dimension $n_R + n_W$, les sommets sont définis par l'intersection de $n_R + n_W$ contraintes. Chaque composante n'apparaît que dans deux contraintes (cf section précédente). Le nombre d'intersections possibles est donc $2^{n_R+n_W}$ et le nombre de sommets lui est inférieur puisqu'on peut obtenir des sommets égaux avec des combinaisons de contraintes différentes quand des paires de contraintes sont équivalentes à une égalité.

Théorème 0.4-6 : le nombre de sommets rationnels d'un polytope de dépendance $D^{W,R}$ est inférieur ou égal à $2^{n_R+n_W}$, ou n_R et n_W sont les nombres de boucles des nids R et W.

Preuve: dans un espace de dimension $n_R + n_W$, les sommets sont définis par l'intersection de $n_R + n_W$ contraintes. Chaque dimension i de w n'apparait que dans deux contraintes parallèles d'intersection vide. Si une composante j de r apparaît aussi dans ces deux contraintes, il faut commencer par la déterminer par d'autres intersection. Sinon, la composante w_i du sommet peut prendre deux valeurs définies par ces deux contraintes.

Les composantes j de r sont toutes définies par l'hypercube d'itérations de$R$. Elles peuvent prendre chacune deux valeurs et les composantes r du sommet, 2^{n_R} valeurs.

Les composantes de w dépendant de r peuvent aussi prendre deux valeurs.

Chaque composante d'un sommet peut prendre deux valeurs. Le nombre de sommets est donc $2^{(n_R+n_W)}$.

QED

Autre preuve: les contraintes non-redondantes sont deux à deux parallèles. A une transformation affine près, elles définissent un hypercube de dimension n_R+n_W. Le nombre de sommets est donc $2^(n_R + n_W)$.

QED

Les bases de l'algorithme de calcul des coordonnées des sommets sont maintenant exposées.

4.5.1 Projection des sommets rationnels sur I^R

Toutes les itérations de R sont effectuées. Les n_R premières coordonnées ont donc pour valeur les bornes de boucles, 0 ou $I^R - 1$. Ceci définit 2^{n_R} premières parties de vecteurs.

Théorème 0.4-7 : Les projections des sommets rationnels du polytope de dépendance sur I^R sont les 2^{n_R} vecteurs obtenus en combinant les bornes de l'ensemble d'itération I^R, i.e. les 2^{n_R} sommets de l'hypercube I^R.

Preuve (trop) rapide: toutes les contraintes sur les coordonnées de w sont des fonctions monotones croissantes des cordonnées de r. Les maximas et les minimas sont atteints sur les sommets du domaines I^R. *Ça ne les empêche pas d'avoir une intersection intermédiaire. Il faut s'appuyer sur la décomposition dimension par dimension...*

Preuve (par l'absurde): supposons qu'il existe un sommet s du polytope de dépendance dont les coordonnées sur I^R n'appartiennent pas aux sommets de cet hypercube. Il existe au moins une coordonnée i de s telle que:

$$0 < s_i < I_i^R - 1$$

Il existe donc par construction une itération r dont la coordonnée r_i est égale à s_i. La contrainte d'appartenance à l'espace d'itération n'est pas suffisante puisqu'elle s'écrit:

$$0 \leq s_i \leq I_i^R - 1$$

Il existe donc une dimension de tableau k tel que ${\Omega_{pav}}_{k,i}^R$ est strictement positif puisque c'est le seul autre type de contraintes (cf. définition du polytope de dépendance). Une des quatre contraintes suivantes est vérifiée par le sommet s_i:

$$\begin{array}{rcl}
0 & = & {\Omega_{pav}}_{k,i}^R r_i + K_k^R - M_k^W + 1 \\
0 & = & {\Omega_{pav}}_{k,i}^R r_i + {\Omega_{fit}}_{k.k}^R (M_k^R - 1) + K_k^R \\
{\Omega_{pav}}_{k,j}^W w_j & = & {\Omega_{pav}}_{k,i}^R r_i + K_k^R - M_k^W + 1 \\
{\Omega_{pav}}_{k,j}^W w_j & = & {\Omega_{pav}}_{k,i}^R r_i + {\Omega_{fit}}_{k.k}^R (M_k^R - 1) + K_k^R
\end{array}$$

Comme r_i n'est pas une valeur extrémale, on peut l'augmenter en r_i' et la première contrainte conduit à un élément non défini puisque ${\Omega_{pav}}_{k,j}^R > 0$:

$$M_k^W - 1 = {\Omega_{pav}}_{k,i}^R r_i + K_k^R < {\Omega_{pav}}_{k,i}^R r_i' + K_k^R$$

On peut aussi la diminuer et la deuxième contrainte conduit à définir un élément de tableau de coordonnées négatives.

Les deuxième et troisième contraintes nécessitent une intersection supplémentaire pour définir un sommet. Cette intersection ne peut pas se faire avec des contraintes ne portant que sur r_i par hypothèse sur la valeur de s_i, $0 \leq s_i \leq I_i^R - 1$.

Il faut donc disposer de contraintes sur w_j. Si on remplace w_j par des valeurs extrémales de l'ensemble des itérations I^W, on retombe dans le cas précédent et on peut montrer qu'on utilise un élément de tableau non défini. Il faut donc disposer d'une autre contrainte sur w_j qui met en jeu une ligne k' du tableau X et une dimension i' de l'ensemble des itérations I^R. L'existence de k' viole une des contraintes APOTRES sur Ω_{pav}^W.

Ou bien il faut disposer d'une autre contrainte sur r_i. Comme on a supposé jusqu'ici qu'une colonne de la matrice Ω_{pav}^R pouvait comporter plusieurs éléments non-nuls, r_i peut être contraint par d'autre dimensions k' du tableau et d'autres dimensions j' de I^W. En éliminant r_i entre les deux hyperplans, on obtient une nouvelle contrainte entre w_j et w_j' qu'on ne peut plus intersecter qu'avec une contrainte de I^W. On va donc toucher des valeurs extrémales dans I^W et on va pouvoir les dépasser en accroissant ou décroissant r_i ce qui conduit encore à un élément de X non défini.

QED

Cette preuve semble fortement appuyée sur la structure de Ω_{pav}^W qui est imposée par l'assignation unique et donc le caractère fonctionnelle de la spécification. Elle s'appuie aussi indirectement sur l'élimination de la fonction modulo dans les expressions d'accès aux tableaux.

Il serait utile de vérifier et de simplifier cette preuve!

4.5.2 Coordonnées des rationnels sommets sur I^W

Théorème 0.4-8 : Les coordonnées des sommets rationnels sur les dimensions relatives à W sont définis par des fonctions des coordonnées des sommets rationnels sur les dimensions relatives à R.

Certaines itérations de W seulement doivent être effectuées avant les itérations de R. Soit i une des dimensions de W pour laquelle on veut calculer les valeurs de cette coordonnée pour tous les sommets. Plusieurs cas se présentent:

- l'indice de boucle w_i correspondant n'est pas utilisé

  $$\forall k {\Omega_{pav}}_{k,i}^W = 0$$

  et une erreur dans les spécifications est détectée parce que le même élément de tableau est défini plusieurs fois;
- sinon il existe au moins une dimension k du tableau X qui est impactée
  - s'il existe deux dimensions impactées, i.e. une colonne de Ω_{pav}^W avec deux éléments non nuls, il y a une erreur de spécification parce que le nid de boucles W ne définit pas un hypercube X
  - sinon, il existe un et un seul k tel que ${\Omega_{pav}}_{k,i}^W \neq 0$
    - si la ligne k de la matrice Ω_{pav}^R est nulle, alors w_i est contraint par:

      $${\Omega_{pav}}_{k,i}^W w_i \geq K_k^R - M_k^W + 1 \ \wedge \ {\Omega_{pav}}_{k,i}^W w_i \leq {\Omega_{fit}}_{k,k}^R (M_k^R - 1) + K_k^R$$

      en supposant que Ω_{fit} est systématiquement définie comme une matrice diagonale. Les deux valeurs possible pour la coordonnée $n_R + i$ des sommets sont les deux valeurs extrêmes de w_i. Cette paire de contrainte est plus stricte que les contraintes définissant l'ensemble des itérations W parce que tous les éléments utilisés sont définis par hypothèse.

- sinon, il existe j tel que ${\Omega_{pav}}^R_{k,j} \neq 0$
  - s'il existe plusieurs j, alors Ω^R_{pav} ne respecte pas les hypothèses
  - sinon, w_i est contraint par:

$$\begin{aligned} &\left({\Omega_{pav}}^W_{k,i} w_i \geq {\Omega_{pav}}^R_{k,j} r_j + K^R_k - M^W_k + 1\right) \\ \wedge \quad &\left({\Omega_{pav}}^W_{k,i} w_i \leq {\Omega_{pav}}^R_{k,j} r_j + {\Omega_{fit}}^R_{k.k}(M^R_k - 1) + K^R_k\right) \end{aligned}$$

    Les deux valeurs possible pour la coordonnée $n_r + i$ des sommets sont les deux valeurs extrêmes de w_i obtenues en remplaçant r_j par sa valeur déjà calculée pour le sommet courant, s_j.

Ce résultat s'obtient sans faire d'hypothèse sur le nombre d'éléments non nuls dans une colonne de Ω^R_{pav}. Le même indice de boucle de lecture peut donc apparaître plusieurs fois dans les expressions d'indices d'accès à un même tableau, contrairement aux hypothèses actuelles.

Cette extension ne s'applique pas directement au calcul des coordonnées des sommets entiers, surtout en présence de partitionnement, parce que le coefficient de w a un impact sur certains coefficient de r. En entier, l'indépendance des coordonnées en r par rapport aux coordonnées en w est mise à mal.

4.6 Prise en compte des tâches atomiques

Les tâches ne comptant qu'une itération doivent être considérées comme corps d'une boucle de profondeur inférieure ou égale à la dimension du tableau défini ($n_W \leq n_X$). Les bornes supérieures des boucles sont nulles. L'ensemble des itérations est l'hypercube de côté 1 et de dimension n_W. La seule itération exécutée est donc l'itération 0.

La présence de bornes de boucles supérieures nulles risque d'induire un manque de contrainte sur le coefficient d'ordonnancement correspondant puisqu'il n'a aucun impact: il est toujours multiplié par zéro. Des contraintes venant d'autres modèles peuvent partiellement combler l'absence totale de contrainte. On risque néanmoins d'explorer un grand nombre de solutions équivalentes pour l'ordonnancement.

Il faut donc ajouter dans la suite de tests de non-régression des applications contenant des bornes supérieures de boucles nulles.

4.7 Ordonnancements des tâches sans entrée ou sans sortie

Les contraintes d'ordonnancement dues aux dépendances sont moins nombreuses. Mais les contraintes de capacités mémoire vont néanmoins favoriser certains ordonnancements.

De telles tâches doivent apparaître dans la suite de tests de non-régression.

4.8 Conclusion sur la construction des sommets rationnels du polyèdre de dépendance

Comme il est probable qu'elle soit équivalente à ce qui précède, il ne semble pas très utile d'entreprendre la correction de la proposition 9.1.4, si ce n'est pour vérifier d'une manière différente que la construction du polyèdre de dépendance est maintenant bien maîtrisée et qu'une partie plus importante que prévu d'APOTRES doit être modifiée.

Les théorèmes développés dans cette section ne sont valables que si l'hypothèse d'assignation unique des éléments de tableau et l'hypothèse de défnion des éléments de tableau utilisés sont toutes les deux vérifiées. Ils supposent aussi que les matrices d'accès sont toutes positives et s'obtiennent comme permutation de matrices diagonales, sauf si on lève cette hypothèse trop forte sur l'accès en lecture défini par Ω^R_{pav}.

Un solveur comme APOTRES doit donc poser explicitement ces conditions et ne pas supposer que la spécification écrite par l'utilisateur les vérifie naturellement.

5 Ordres partiels sur les sommets et minimisation du nombre des contraintes induites pour l'ordonnancement

Le nombre de sommets du polytope de dépendance est une fonction exponentielle de la somme des profondeurs des deux nids de boucles. Certains de ces sommets sont redondants sous des hypothèses faites pour d'autres modèles d'APOTRES.

En toute rigueur, il est interdit d'affaiblir le modèle de dépendance en prenant en compte d'autres modèles comme l'architecture ou l'ordonnancement. En pratique, il peut être important de réduire le nombre de contraintes sur l'ordonnancement dues aux dépendances, en particulier quand on essaie de calcuelr manuellement des ordonnancements.

5.1 Impact de l'ordre lexicographique sur les itérations (ordre presque total)

On suppose que tous les nids de boucle ont un ordonnancement compatible avec l'ordre lexicographique, $r' \succeq r \Rightarrow d^R(r') \geq d^R(r)$, c'est-à-dire que la fonction d'ordonnancement est croissante.

Théorème 0.5-1 : Soit un sommet $s = (r, w)$ et un sommet $s' = (r', w')$. Si l'itération r' est postérieure ou égale à l'itération r, $r' \succeq r$, et si l'itération w' est antérieure ou égale à l'itération w, alors le sommet s' n'apporte pas d'information par rapport à s.

Preuve: Soit d^W et d^R, comme *date*, les deux fonctions d'ordonnancement associées à W et R..

Le sommet s est équivalent à $d^R(r) > d^W(w)$. Comme par hypothèse $d^R(r') \geq d^r(r)$ et $d^W(w) \geq d^W(w')$, on a $d^R(r') \geq d^R(r) > d^W(w) \geq d^W(w')$, i.e. $d^R(r') > d^W(w')$. Le sommet s et les ordres lexicographiques impliquent donc le sommet s'.

QED

Dans le cas particulier où le point $(0, I^W - 1)$ appartient à l'ensemble des dépendances, ce seul sommet suffit à caractériser toutes les dépendance possibles (r', w') puisque $r' \succeq 0$ et $w' \preceq I^W - 1$. C'est le cas quand l'espace des dépendance est un hyper-rectangle parce que seul les motifs interviennent dans les fonctions d'accès.

Le nombre de contraintes d'ordonnancement utiles est donc largement inférieure à $2^{n_W + n_R}$. On pourrait essayer d'en profiter lors de la construction des sommets ou bien se contenter de les filtrer ensuite avant de les utiliser.

Théorème 0.5-2 : Pour chaque sommet r de I^R, il existe une seule dernière itération w dans I^W permettant de construire un sommet (r, w), obtenue en choisissant pour chaque coordonnee de w sa valeur maximale en fonction de r. On a donc au maximum 2^{n_R} sommets utiles pour l'ordonnancement.

Par exemple, si le nid R utilisant le tableau X ne comporte qu'une boucle, on aura besoin au plus de deux sommets, un qui servira a définir γ et un qui servira à définir δ_1^R.

5.2 Impact des coefficients positifs des ordonnancements (ordre partiel)

Même si l'ordre lexicographique n'est pas imposé, il est dit que les coefficients δ sont tous positifs. La fonction d'ordonnancement est donc une fonction croissante par rapport à toutes les coordonnées des vecteurs d'itérations.

Théorème 0.5-3 : Soit un sommet $s = (r, w)$ et un sommet $s' = (r', w')$. Si l'itération r' est plus grande ou égale à l'itération r, $r' \geq r$, et si l'itération w' est inférieure ou égale à l'itération w, $w' \leq w$, alors le sommet s' est valide pour toute paire d'ordonnancements croissants (d^R, d^W) valide par rapport à s.

$$\forall (d^R, d^W) \in \mathbb{N}^{n_R+1} \times \mathbb{N}^{n_W+1} d^W(w) \leq d^R(r) \quad \wedge \quad r' \geq r \quad \wedge w' \leq w \Longrightarrow d^W(w') \leq d^R(r')$$

Preuve: Par définition d'un ordonnancement croissant, $r' \geq r$ implique $d^R(r') \geq d^R(r)$ et $w' \leq w$ implique $d^W(w') \leq d^W(w)$. On a donc:

$$d^W(w') \leq d^W(w) \quad \wedge \quad d^W(w) \leq d^R(r) \quad \wedge \quad d^R(r) \leq d^R(r')$$

ce qui implique la conclusion $d^W(w') \leq d^R(r')$.
QED
Même sans prendre en compte l'ordre lexicographique, certains sommets du polytope de dépendance sont inutiles pour poser les contraintes de dépendances. Mais il n'est pas possible d'associer à chaque sommet de I^R une seule itération maximale w parce que l'ordre est partiel. Le nombre de sommets utiles à l'ordonnancement n'est donc pas (?) au plus de 2^{n_R} sous l'hypothèse de positivité des coefficients d'ordonnancement.
A REVOIR. THEOREME A ECRIRE ET A PROUVER. ON S'APPUIE TROP SUR LE CAS DES DEPENDANCES 2-D OU L'ORDRE SUR LES COMPOSANTS EST EGAL A L'ORDRE LEXICOGRAPHIQUE PUISQUE CHAQUE ESPACE D'ITERATION EST MONODIMENSIONNEL.

Théorème 0.5-4 : Si l'espace des dépendances est bidimensionnel, le nombre de sommets utiles est deux et non quatre si les ordonnancements sont à coefficients positifs.

Preuve: l'ordre sur les composants est égal à l'ordre lexicographique. Le théorème portant sur l'ordre lexicographique s'applique.

5.3 Impact du non-couplage

Dans le cas où une dimension i de r n'est couplée à aucune dimension j de w via une dimension k du tableau X et les matrices Ω^W_{pav} et Ω^R_{pav}, le nombre de sommets est divisés par deux en ne prenant en compte que les sommets $g = (r, w)$ où $r_i = 0$ puisqu'ils dominent tous les sommets où $r_i \neq 0$.

5.4 Autres possibilités?

Plus généralement, une fonction d'ordonnancement du nid de boucles R est définie par n_R+1 constantes, n_R variables δ^R et une variable γ^R.

L'objet de cette section est de voir s'il est possible de réduire le nombre de contraintes d'ordonnancement, i.e. le nombre de sommets utiles à l'ordonnancement, à une fonction linéaire du nombre de boucles, par exemple $n_R + 1$. On suppose que les ordonnancements sont tous compatibles avec l'ordre lexicographique.

Si on choisit le point $r = 0$, qui appartient toujours à I^R par hypothèse sur les bornes de boucle et qui est toujours une composante d'au moins un sommet puisqu'elle ne peut pas être obtenue comme combinaison linéaire de points positifs, on obtient une contrainte sur γ^R et sur l'ordonnancement de W fondée sur le sommet $g_0 = (0, w_0)$ qui est le plus grand sommet associé à $r = 0$ pour l'ordre lexicographique.

$$\delta^W w_0 \leq \gamma^R$$

Pour simplifier les calculs, on a pris $\gamma^W = 0$, ce qui peut toujours être obtenu par une translation.

Si on choisit ensuite les autres sommets r de I^R tel que toutes les coordonnées de r soit nulles sauf une, la i-ème, on obtient chaque fois une contrainte portant uniquement sur δ_i^R et γ^R et sur l'ordonnancement de W. En appelant r_i le vecteur dont seule la i-ème composante est non nulle, et en posant $g_i = (r_i, w_i)$ où w_i est le plus grand vecteur associé à r_i pour l'ordre lexicographique, on a:

$$\delta^W w_i \leq \delta^R r_i + \gamma^R$$

Il faut noter que le nombre de vecteurs r_i, le nombre de vecteurs w_i et donc le nombre de vecteurs $g_i = (r_i, w_i)$ sont tous les trois égaux à n_R.

Théorème 0.5-5 : Les w_i maximaux sont liés au r_i par une fonction affine f dont la valeur à l'origine est w_0.

Preuve: d'après le théorème xyz sur l'analyse des contraintes de dépendances, chaque composante j de w_i est soit une constante, et donc une fonction affine de r_i, soit elle est définie par:

$$(w_i)_j = \frac{\Omega_{pav}{}_{k,j'}^{R}(r_i)'_j + \Omega_{fit}{}_{k.k}^{R}(M_k^R - 1) + K_k^R}{\Omega_{pav}{}_{k,j}^{W}}$$

qui définit sa valeur maximale. Comme chaque composante de w_i est définie par une fonction affine, le vecteur w_i est défini par une fonction affine.

QED

Soit f cette fonction affine. On peut donc écrire $f(r + r') = f(r) - w_0 + f(r') - w_0 + w_0$.

L'itération maximale w associée à l'itération $r_i + r_j$ est donc $w_i + w_j - w_0$. On voudrait pouvoir prouver, entre autres, que l'ordonnancement pour une combinaison de sommets:

$$\delta^W (w_i + w_j - w_0) \leq \delta^R (r_i + r_j) + \gamma^R$$

est valable sachant qu'il est valable sur les sommets g_i, g_j et g_0:

$$\delta^W w_0 \leq \gamma^R \quad \wedge \quad \delta^W w_i \leq \delta^R r_i + \gamma^R \quad \wedge \quad \delta^W w_j \leq \delta^R r_j + \gamma^R$$

Conjecture fausse: Un ordonnancement valable pour les n_R sommets g_i et pour le sommet g_0 est valable pour tous les sommets du polyèdre de dépendance et donc pour tous les vecteurs de ce polyèdre.

Preuve de l'erreur: si ce théorème est vrai, il est aussi vrai pour le cas particulier construit à partir de r_i et r_j. L'itération $r_i + r_j$ doit dépendre de l'itération $w_i + w_j - w_0$ puisque la dernière itération de définition est une fonction affine de l'itération de lecture (cf. théorème précédent). Ceci s'écrit:

$$\left.\begin{array}{l} \delta^W w_0 \leq \gamma^R \\ \delta^W w_i \leq \delta^R r_i + \gamma^R \\ \delta^W w_j \leq \delta^R r_j + \gamma^R \end{array}\right\} \Longrightarrow \delta^W (w_i + w_j - w_0) \leq \delta^R (r_i + r_j) + \gamma^R$$

où la proposition impliquée est une formalisation de la conjecture qui dit que la contrainte d'ordonnancement est valable pour tous les sommets du polyèdre de dépendance. En effet, l'itération $r_i + r_j$ appartient par construction à l'ensemble d'itération I^R et $w_i + w_j - w_0$ est la plus grande itération dont elle dépend.

C'est équivalent à dire que les trois premières contraintes et la négation de la dernière ne peuvent pas être vraies simultanément, ou bien à dire que l'ensemble ainsi défini doit être vide puisque $A \Rightarrow B \equiv \bar{A} \vee B \equiv \overline{A \wedge \bar{B}}$:

$$\left\{\begin{array}{l} \delta^W w_0 \leq \gamma^R \\ \delta^W w_i \leq \delta^R r_i + \gamma^R \\ \delta^W w_j \leq \delta^R r_j + \gamma^R \\ \delta^R (r_i + r_j) + \gamma^R \leq \delta^W (w_i + w_j - w_0) - 1 \end{array}\right.$$

La constante -1 est utilisée pour supprimer la comparaison stricte parce que toutes les dates considérées sont entières. En additionnant les trois dernières contraintes, on obtient:

$$\left\{\begin{array}{l} \delta^W w_0 \leq \gamma^R \\ \delta^W w_i + \delta^W w_j + \delta^R (r_i + r_j) + \gamma^R \leq \delta^W (w_i + w_j - w_0) - 1 + \delta^R r_i + \delta^R r_j + 2\gamma^R \end{array}\right.$$

soit:

$$\left\{\begin{array}{l} \delta^W w_0 \leq \gamma^R \\ 0 \leq -\delta^W w_0 - 1 + \gamma^R \end{array}\right.$$

ou bien:

$$\left\{\begin{array}{l} \delta^W w_0 \leq \gamma^R \\ \delta^W w_0 \leq \gamma^R - 1 \end{array}\right.$$

Cet ensemble ne peut être vide tout en correspondant à un ordonnancement valide que si:

$$\gamma^R = \delta^W w_0$$

ce qui viole la deuxième contrainte. On ne peut donc envisager de se passer de certains sommets que si on ordonnance la première itération de la tâche R au plus tôt: $\gamma^R = \delta^W w_0$.

QED

Donc, dans le cas général, on ne peut pas se passer des autres sommets et ne considérer que $n_R + 1$ sommets, sauf, peut-être, si on ordonnance la première itération de R au plus tôt en posant $\gamma^R = \delta^W w_0$.

Intuitivement, on peut comprendre le phénomène dans un espace de dépendance tridimensionnel formé d'une boucle d'écriture et de deux boucles de lecture. La paire des ordonnancement pour W et R est équivalente à une forme affine donnant le délai. Si cette forme affine est plaquée au polyèdre de dépendance, sa pente va être forcée par les premiers sommets et forcément respecter leurs sommes.

Si cette forme affine définit un grand délai pour l'itération 0, il va être possible d'enchaîner ensuite plus rapidement les itérations de r mais la plus difficile à placer va alors être la dernière, $r_1 + r_2$, ou dans le cas général, $\sum_{i\in[1,n_R]} r_i$.

ISABELLE: tu fais un dessin 2 ou 3D?

Dans le cas où le délai est encore plus grand, par exemple si γ^R est plus grand que la dernière itération utile, $\sum_{i\in[1,n_R]} \delta^W w_i$, ou si γ^R est plus grand que la dernière itération de définition, $\delta^W I^W$, il ne reste plus que les contraintes d'ordre lexico-positif et donc plus qu'un unique ordonnancement possible pour R. Le parcours des coefficients δ^R est alors devenu inutile.

ISABELLE: tu fais un dessin 2 ou 3D?

Théorème 0.5-6 : Un ordonnancement vérifiant $\gamma^R = \delta^W w_0$ et valide pour les n_R sommets g_i de $\bar{D}^{W,R}$ est un ordonnancement valide pour $\bar{D}^{W,R}$.

Preuve: les composantes w des sommets dominants du polyèdre de dépendance sont des fonctions affines des composantes r (cf. Théorème???). Les composantes r de n'importe quel sommet appartiennent à l'hypercube défini par 0 et les r_i et sont obtenues par:

$$r = \sum_{i\in[1,n_R]} \lambda_i r_i$$

sous la condition:

$$\forall i \in [1, n_R] \quad \lambda_i \in [0, 1]$$

Note: Nicolas suggère de prendre $\lambda_i \in \{0, 1\}$. Je ne sais pas trop si je veux faire la preuve pour n'importe quelle itération r ou seulement pour l'ensemble des sommets du polyèdre de dépendance. A REVOIR

Le point w associé à r est donc:

$$w = w_0 + \sum_{i\in[1,n_R]} \lambda_i (w_i - w_0)$$

Par hypothèse, comme les n_R vecteurs $g_i = (r_i, w_i)$ sont des dépendances satisfaites et comme les λ_i sont positifs:

$$\lambda_i \delta^W w_i \leq \lambda_i \delta^R r_i + \lambda_i \gamma^R$$

Par hypothèse, $\gamma^R = \delta^W w_0$. On peut donc écrire:

$$\sum_{i\in[1,n_R]} \lambda_i \delta^W w_0 = \sum_{i\in[1,n_R]} \lambda_i \gamma^R$$

En soustrayant cette équation de la somme sur i des inéquations précédentes, on obtient:

$$\sum_{i\in[1,n_R]} \lambda_i \delta^W (w_i - w_0) \leq \sum_{i\in[1,n_R]} \lambda_i \delta^R r_i$$

soit par définition de r et de w, l'inégalité:

$$\delta^W (w - w_0) \leq \delta^R r$$

auquel on ajoute l'égalité choisit comme hypothèse, $\gamma^R = \delta^W w_0$, pour obtenir:

$$\delta^W w \leq \delta^R r + \gamma^R$$

Cette inégalité signifie par définition que l'ordonnancement est compatible avec le sommet (r, w).

QED

Note: il doit y avoir un bug dans la preuve puisque l'hypothèse $\forall i \in [1, n_R] \quad \lambda_i \in [0, 1]$ n'est pas utilisée. Mais peut-être pas: on semble se ramener au cas d'un cône de dépendance. Les sommets considérés sont 0 et r_i. On semble pouvoir traiter les g_i comme des rayons plutôt que comme des sommets.

La conclusion pratique est qu'APOTRES devrait toujours trouver très rapidement un ordonnancement valide.

Note: On ne peut pas appliquer la même approche pour dériver l'ordonnancement de W à partir de l'ordonnancement de R. R ne dépend que d'un W alors que W peut fournir des données à plusieurs nids de boucles. De plus, par construction, on sait que $r = 0$ est une composante de sommets de $D^{W,R}$, mais ce n'est pas le cas de $w = 0$.

5.5 Impact négatif de la contrainte actuelle de sérialisation

Notons enfin que l'architecture SIMD impose des contraintes supplémentaires sur les ordonnancements qui peuvent prendre le pas sur les contraintes provenant des dépendances. Le parallélisme est traité à un niveau moyen ou fin par le modèle de partitionnement et l'ordonnancement doit être séquentiel. Comme toutes les itérations d'une tâche peuvent être exécutées en parallèle, *i.e.* dans n'importe quel ordre, on choisit arbitrairement d'imposer l'ordre lexicographique.

Pour toutes les tâches R (ou W!), on impose que:

$$\gamma^R \geq 0$$

$$\forall i \in [1, n_R] \quad \delta_i^R \geq 1$$

$$\forall i \in [1, n_R[\quad \delta_i^R \geq \sum_{j \in]i, n_R]} \delta_j^R (I_j^R - 1)$$

Note pour Nicolas: j'avais ferme le premier interval pour obtenir l'ensemble vide pour la sommation et la contrainte précédente $\delta_{n_R}^R \geq 1$, mais on obtient en fait seulement $\delta_{n_R}^R \geq 0$. Il faut donc maintenir la contrainte precedente et celle-ci est redondante.

La troisième contrainte implique la seconde *(faux d'après la remarque ci-dessus)* ainsi que l'ordre lexicographique sur les itérations qui a été utilisé pour réduire le nombre de sommets à prendre en compte pour poser les contraintes d'ordonnancement.

Prenons l'exemple:

```
DO i1 = 0, 9
  DO j1 = 0, 9
    X[i1, j1] = ...

DO i2 = 0, 9
  DO j2 = 0, 9
    ... = X[j2, i2]
```

Les dépendances imposent i1==j2 et j1==i2. Les sommets sont donc de la forme (i2, j2, j2, i2) avec les quatre combinaisons possibles pour les valeurs du couple (i2,j2):

- (0,0,0,0): $\gamma^W \leq \gamma^R$ ce qui permet de prendre ces deux constantes égales à zéro;
- (0,9,9,0): $9\delta_1^W \leq 9\delta_2^R$ ou $\delta_1^W \leq \delta_2^R$
- (9,0,0,9): $9\delta_2^W \leq 9\delta_1^R$ ou $\delta_2^W \leq \delta_1^R$
- (9,9,9,9): est redondante avec les deux précédentes parce que les deux constantes sont nulles.

Les contraintes lexicographiques sont $9\delta_2^W < \delta_1^W$ et $9\delta_2^R < \delta_1^R$ ce qui donne pour des coefficients d'ordonnancement strictement positifs:

$$0 < 81\delta_2^W < 9\delta_1^W \leq 9\delta_2^R < \delta_1^R$$

dont la plus petite solution est:

$$\delta_2^W = 1, \delta_1^W = 10, \delta_2^R = 10, \delta_1^R = 91$$

Les deux premières itérations de chaque boucle se suivent immédiatement, mais la seconde itération de la deuxième boucle ne peut pas se faire avant la date 10 parce que les dimensions ne sont pas permutables dans le cadre actuel.

Les permutations de boucles sont cependant prises en compte par APOTRES par un autre mécanisme. Voici le meilleur ordonnancement proposé par APOTRES:

Contribution NICOLAS MUSEUX

5.6 Conclusion sur les sommets rationnels redondants

D'un pur point de vue ordonnancement et PLC, aucun sommet rationnel du polytope de dépendance n'est redondant. La contrainte actuelle de sérialisation des itérations induit d'ailleurs des problèmes parce qu'elle ne permet pas de traiter l'échange de boucles comme un problème d'ordonnancement.

Les résultats de cette section peuvent être utilisé en résolution manuelle, ou éventuellement en résolution automatique pour gagner du temps. Mais attention à l'indépendance des modèles et à leur réutilisation pour d'autres machines cibles.

6 Proposition 9.1.4 du rapport DRET de 1996

La proposition 9.1.4 consiste à reconstruire le polyèdre de dépendance à partir d'un petit nombre de projections sur des plans, puis à reconstruire les sommets du polyèdre de dépendance à partir des sommets des projections. Sa preuve a été considérée comme trop simple pour être donnée dans le rapport DRET. La procédure de reconstruction des sommets n'a pas été définie. Il suffit pourtant de considérer le cas plus simple de la reconstruction d'un polygone à partir de ses projections sur des droites pour se rendre compte que le problème n'est pas évident.

Nous montrons d'abord que la proposition est erronée, puis nous donnons de nouvelles hypothèses sous lesquelles la décomposition proposée est correcte, enfin nous l'étendons pour la rendre valable sous les hypothèses APOTRES mais cette extension semble contre-productive pour la reconstruction des sommets.

6.1 La proposition 9.1.4 est erronnée

La proposition 9.1.4 établit une égalité générale entre un sous-ensemble d'un espace de dimension $n_R + n_W$ et le produit cartésien de n_X sous-ensembles d'espace de dimension 2:

Définition: La projection $D_k^{W,R}$ de $D^{W,R}$ est définie par:

$$D_k^{W,R} = \left\{(x,y)|\exists!i\exists!j\exists(r,w) \in D^{W,R} \quad x = w_i \quad \wedge \quad y = r_j \quad \wedge \quad \Omega_{pav}{}_{k,i}^{W} \neq 0 \quad \wedge \quad \Omega_{pav}{}_{k,j}^{R} \neq 0\right\}$$

Cet ensemble risque d'être vide pour certaines valeurs de k puisque par hypothèse il existe au plus un i et au plus un j. Ils n'existent pas forcément.

Proposition 9.1.4: L'ensemble des dépendances $D^{W,R}$ est égal au produit cartésien des n_X projections $D_k^{W,R}$ de $D^{W,R}$ sur des plans:

$$D^{W,R} = \otimes_{k\in[1,n_X]} D_k^{W,R}$$

Au mieux, du fait de l'égalité des dimensions, elle ne peut être vraie que si $n_R + n_W = 2n_X$. Comme $n_W \leq n_X$ du fait de l'assignation unique (cf. Théorème 4), il faudrait que $n_R \geq n_X$.

Le fait que certaines projections, telles qu'elles sont définies, puissent être vides invalide la reconstruction parce que l'ensemble des dépendances n'est jamais vide sous les hypothèses faites alors que le produit cartésien avec un ensemble vide produit un ensemble vide.

Le simple remplacement du produit cartésien par un produit plus général permettant d'assimiler certaines dimensions ne suffit pas. On aurait alors $n_R + n_W \leq 2n_X$. Le principe même de la décomposition de l'espace des dépendances dimension de tableau par dimension de tableau est mis à mal, même si les dépendances ne mettant pas en jeu une dimension de tableau, une dimension de l'espace de définition et une dimension de l'espace d'utilisation ne peuvent être que triviales puisqu'elles se ramènent à des produits d'intervalles.

De plus, la proposition 9.1.4 suppose l'existence d'une procédure de reconstruction des sommets à partir des sommets des projections.

6.2 Réécriture de la proposition 9.1.4

La proposition 9.1.4 est vraie quand $n_W = n_R = n_X$ et quand les rangs de Ω_{pav}^R et Ω_{pav}^W sont égaux à n_X.

Preuve: les contraintes peuvent être partitionnées en k ensembles disjoints de dimension 2. Les coordonnées des sommets des projections et les coordonnées des sommets du polyèdre de dépendances se calculent avec les mêmes intersections.

Théorème 0.6-1 : Si $n_W = n_R = n_X$ et si le rang de Ω_{pav}^R est égal à n_X, les sommets du polyèdre de dépendance sont obtenus en combinant de toutes les manières possibles les coordonnées des sommets des projections.

Preuve (?): Il existe n_X colonnes indépendantes dans Ω_{pav}^R puisque son rang est n_X. Chaque colonne a par construction au plus un élément non nul. On sait associer à chaque ligne k une unique colonne i_k. Le rang de Ω_{pav}^W est n_W par définition. On sait aussi associer à chaque ligne k une unique colonne j_k.

Comme chaque colonne est associée à au plus une ligne par hypothèse sur les matrices d'accès, on peut donc traiter indépendemment les n_X paires de dimensions (i_k, j_k). Les contraintes de dépendances sont exactement partitionnées en n_X ensembles de contraintes disjointes puisque pour tout $k' \neq k$, $i_{k'} \neq i_k$ et $j_{k'} \neq j_k$.

QED

Le nombre de sommets rationnels du polyèdre de dépendance définis par la proposition 9.1.4 est 4^{n_X} au plus, soit 2^{2n_X} au plus.

Le nombre de sommets du polytope de dépendance utiles pour poser les contraintes d'ordonnancement est $2^{n_X} = 2^{n_R}$ parce que pour chaque coordonnée r_i on peut maximiser la coordonnée w_j qui lui correspond via la dimension k du tableau X.

Note: le nombre de sommets rationnels du polytope de dépendance n'est néanmoins pas toujours 4 ou 6 comme dans la version actuelle d'APOTRES (cf. section???).

6.3 Remplacement de la proposition 9.1.4

Comme $n_W + n_R$ peut être beaucoup plus grand que n_X puisque n_R n'est pas contraint par n_X, il faut augmenter le nombre de projections à partir des sommets desquelles on reconstruit les sommets du polyèdre de dépendance.

Comme on ne bénéficie plus d'un partitionnement des dimensions, la reconstruction des sommets du polytope des dépendances à partir des sommets de ses projections doit être explicitée.

On définit une famille de projection 2-D de $\bar{D}^{W,R}$ sur une dimension de W et une dimension de R:

$$\bar{D}^{W,R}_{k,l} = \left\{(x,y) \in \mathbb{N}^2 | \exists r \exists w t.q. (r,w) \in D^{W,R} \quad \wedge \quad x = r_k \quad \wedge \quad y = w_l\right\}$$

Théorème 0.6-2 : L'ensemble $D^{W,R}$ est égal à l'ensemble reconstruit à partir de toutes ses projections sur les plans contenant un vecteur de base de I^R et un vecteur de base de I^W.

$$\bar{D}^{W,R} = \left\{(r,w) \in \mathbb{N}^{n_R} \times \mathbb{N}^{n_W} | \forall k \in [1,n_R] \forall l \in [1,n_W] \quad (r_k, w_l) \in \bar{D}^{W,R}_{k,l}\right\}$$

Preuve: par définition de $\bar{D}^{W,R}_{k,l}$, tout élément de $\bar{D}^{W,R}$ vérifie l'ensemble des contraintes de la partie droite. $\bar{D}^{W,R}$ est inclu dans l'ensemble reconstruit à partir des projections.

D'autre part, toute contrainte apparaissant dans le définition de $D^{\bar{W},R}$ fait apparaître au plus deux variables, correspondant à une dimension k du tableau, i de r et j de w. Par définition de la projection, cette contrainte est conservée inchangée dans $D^{W,R}_{i,j}$. Les contraintes portant uniquement sur r ou w ne concernent qu'une coordonnée. Elles sont répliquées dans de nombreuses projections. Toutes les contraintes de $\bar{D}^{W,R}$ étant répliquées dans l'ensemble de droite, il ne peut être qu'inclu dans $\bar{D}^{W,R}$.

Les deux ensembles sont donc égaux.

QED

Note: on n'utilise pas toutes les projections 2-D possibles. Les projections sur des plans de $\mathbb{N}^{n_R}$ ou sur des plans de $\mathbb{N}^{n_W}$ sont inutiles.

6.4 Conclusion sur la proposition 9.1.4

La nouvelle reconstruction est correcte, mais elle fait intervenir énormément de contraintes redondantes. Pourquoi devrait-elle conduire à une reconstruction efficace des sommets du polytope de dépendance? Est-il utile de rechercher un algorithme de reconstruction? Un algorithme de reconstruction efficace peut-il faire mieux que l'algorithme de calcul des sommets qui a déjà été proposé?

7 Prise en compte du partitionnement

Du fait de la présence des motifs, le partitionnement pourrait ne pas modifier l'analyse. Les quantificateurs supplémentaires introduits dans la définition des dépendances peuvent éventuellement être pris en compte via un motif modifié, éventuellement fermé par convexité.

Nous vérifions dans cette section que ce n'est vrai que pour la partitionnement de la tâche de définition W et non pour la tâche d'utilisation R.

Plutôt que de repartir de la définition des dépendances, on utilise les ensembles d'éléments utilisés $X^R(r)$ ou définis $X^W(w)$ par une itération.

7.1 Rappel sur le partitionnement

Le partitionnement consiste à faire un changement de repères définis par:

$$i = LPc + Lp + l$$

où i est une itération comme r ou w, c est un nouveau vecteur d'itération, p est un vecteur dénotant un processeur et l est un vecteur dénotant une itération particulière à l'intérieur d'un hypercube d'itération L. P et L sont des matrices diagonales, qui peuvent donc aussi être vus comme des vecteurs. On a de plus:

$$p \in [0, P[= [0, P-1] \quad \wedge \quad l \in [0, L[= [0, L-1]$$

Du fait de la présence des motifs, le partitionnement pourrait ne pas modifier l'analyse. Les quantificateurs supplémentaires introduits dans la définition des dépendances pourraient être pris en compte via un motif modifié, éventuellement fermé par convexité.

7.2 Ensemble des éléments définis par une itération

L'ensemble des éléments définis par une itération c^W pour un partitionnement (P, L) est donc induit par l'ensemble des itérations w de composante c^W:

$$w(c^W) = \left\{ \begin{array}{cc} & w \in \mathbb{N}^{n_W} | \exists p \in [0, P-1] \exists l \in [0, L-1] \\ \text{t.q.} & w = LPc^W + Lp + l \quad \wedge \quad w \in [0, I^W - 1] \end{array} \right\}$$

où w est surchargée en tant qu'élément et en tant que fonction.

L'ensemble des éléments de tableau définis par l'itération de bloc c^W est donc:

$$X_c^W(c^W) = \bigcup_{w \in w(c^W)} X_i^W(w)$$

qui peut se réécrire:

$$X_c^W(c^W) = \left\{ \phi \in \mathbb{N}^{n_X} | \exists w \in w(c^W) \phi \in X_i^W(w) \right\}$$

puis, en reprenant la définition de $w(c^W)$:

$$X_c^W(c^W) = \left\{ \begin{array}{cc} & \phi \in \mathbb{N}^{n_X} | \exists p \in [0, P-1] \exists l \in [0, L-1] \\ \text{t.q.} & w = LPc^W + Lp + l \quad \wedge \quad w \in [0, I^W - 1] \quad \wedge \quad \phi \in X_i^W(w) \end{array} \right\}$$

puis celle de $X_i^W(w) = \left\{ \phi | \exists m \in M^W, \quad w \in I^W \quad \phi = \Omega_{pav}^W w + m \right\}$:

$$X_c^W(c^W) = \left\{ \begin{array}{c} \phi \in \mathbb{N}^{nx} | \exists p \in [0, P-1] \exists l \in [0, L-1] \exists m \in [0, M^W - 1] \\ \text{t.q.} \quad w = LPc^W + Lp + l \quad \wedge \quad w \in [0, I^W - 1] \quad \wedge \quad \phi = \Omega_{pav}^W w + m \end{array} \right\}$$

On peut éliminer w:

$$X_c^W(c^W) = \left\{ \begin{array}{ll} & \phi \in \mathbb{N}^{nx} | \exists p \in [0, P-1] \exists l \in [0, L-1] \exists m \in [0, M^W - 1] \\ \text{t.q.} & LPc^W + Lp + l \in [0, I^W - 1] \\ \wedge & \phi = \Omega_{pav}^W LPc^W + \Omega_{pav}^W Lp + \Omega_{pav}^W l + m \end{array} \right\}$$

Pour se ramener à la définition de $X_i^W(w)$, il faut accepter une légère sur-approximation de $X_c^W(c^W)$ permettant de négliger des contraintes sur p et l pouvant exister lors des dernières itérations. Ces contraintes découlent du reste de la division de $I^W - 1$ par PL. Ajouter ces éléments à l'ensemble des itérations I^W n'a pas en soi d'impact puisqu'on suppose que tous les éléments lus ultérieurement sont définis. En définir en plus ne change pas les dépendances.

$$\bar{X}_c^W(c^W) = \left\{ \begin{array}{c} \phi \in \mathbb{N}^{nx} | \exists p \in [0, P-1] \exists l \in [0, L-1] \exists m \in [0, M^W - 1] \\ \text{t.q.} \quad c^W \in \left[0, \left\lfloor \dfrac{I^W - 1}{PL} \right\rfloor\right] \quad \wedge \quad \phi = \Omega_{pav}^W LPc^W + \Omega_{pav}^W Lp + \Omega_{pav}^W l + m \end{array} \right\}$$

Il faut ensuite définir un nouveau motif m' pour prendre en compte en une seule fois les trois dernier termes:

$$m' = \Omega_{pav}^W Lp + \Omega_{pav}^W l + m$$

Une ligne k de Ω_{pav}^W contient au plus un élément non nul, dont la valeur est alors M_k^W. On a donc:

$$\begin{aligned} m' &\in \left[0, \Omega_{pav}^W L(P-1) + \Omega_{pav}^W (L-1) + M^w - 1\right] \\ &= \left[0, \Omega_{pav}^W LP - \Omega_{pav}^W + \Omega_{pav}^W L - \Omega_{pav}^W + M^w - 1\right] \\ &= \left[0, \Omega_{pav}^W LP - 1\right] \end{aligned}$$

On obtient donc pour $\bar{X}_c^W(c^W)$ une formulaire similaire à celle de $X_i^W(w)$:

$$\bar{X}_c^W(c^W) = \left\{ \begin{array}{c} \phi \in \mathbb{N}^{nx} | \exists m' \in [0, \Omega_{pav}^W LP - 1] \\ \text{t.q.} \quad c^W \in \left[0, \left\lfloor \dfrac{I^W - 1}{PL} \right\rfloor\right] \quad \wedge \quad \phi = \Omega_{pav}^W LPc^W + m' \end{array} \right\}$$

avec un ensemble d'itération réduit par le partitionnement, un motif augmenté par le partitionnement et une matrice d'accès multipliée par les deux matrices de partitionnement, L et P.

Pour prendre en compte le fait que des nids de boucles différents sont partitionnés différemment, il faut réécrire la formule précédente en indexant les matrices de partitionnement:

$$\bar{X}_c^W(c^W) = \left\{ \begin{array}{c} \phi \in \mathbb{N}^{nx} | \exists m' \in [0, \Omega_{pav}^W L^W P^W - 1] \\ \text{t.q.} \quad c^W \in \left[0, \left\lfloor \dfrac{I^W - 1}{L^W P^W} \right\rfloor\right] \quad \wedge \quad \phi = \Omega_{pav}^W L^W P^W c^W + m' \end{array} \right\}$$

Note: est-ce que les motifs sont censés tomber juste sur les dimensions de tableaux? La réponse est probablement: oui, implicitement, parce que les dimensions de tableaux ne sont pas déclarées dans APOTRES. NICOLAS MUSEUX? Dans le cas de la tâche de définition, ça n'a pas d'importance. Dans le cas de la tâche d'utilisation, ça peut ajouter des itérations.

7.3 Ensemble des éléments lus par une itération

L'ensemble des éléments lus est plus difficile à caractériser que l'ensemble des éléments écrits parce qu'on prend l'image du motif et non le motif et parce que la matrice d'accès n'est pas conditionnée par la taille du motif. De plus, on ne peut pas surapproximer l'ensemble des éléments lus sans prendre le risque d'augmenter les dépendances.

Le partitionnement de R est supposé calé à l'origine. On sait donc que la plus petite itération d'un bloc contenant au moins une itération de l'ensemble des itérations appartient toujours à cet ensemble:

$$c^R \in \left[0, \left\lfloor \frac{I^R - 1}{L^R P^R} \right\rfloor\right] \Longrightarrow L^R P^R c^R \in \left[0, I^R - 1\right]$$

Ceci n'est malheureusement faux pour les dernières itérations d'un bloc $P^R L^R$ et, comme la taille de ce bloc est ajustable pendant la résolution, il n'est pas possible de supposer que $P^r L^R$ divise exactement I^R, à moins d'ajouter cette contrainte dans un autre modèle, comme le modèle d'architecture où elle est appelée *contrainte SIMD*.

Supposons que les deux nids de boucles donnés ci-dessous:

```
DO i1 = 0, 6
  X[i1] = ...
DO i2 = 0, 6
  ... = X[i2]
```

soit partitionnées respectivement en blocs de 7 et de 4 éléments pour un unique processeur:

```
DO c1 = 0, 0
  DO l1 = 0, 6
    X[7*c1+l1] = ...
DO c2 = 0, 1
  DO l2 = 0, 3
    ... = X[4*c2+l2]
```

Faute de garde sur la borne supérieure de l2, `min(3, 6-4*c2)`, l'itération 1 de la boucle c2 va sembler dépendre de l'itération 1 de la boucle c1 à cause de l'élément non-défini `X[7]`.

En conclusion, et contrairement à ce qui a été fait dans le rapport DRET et la thèse de Christophe Guettier, il ne semble pas souhaitable de calculer les sommets directement dans les coordonnées c. Il vaudrait mieux les calculer dans les coordonnées initiales, puis faire un changement de coordonnées.

À tout le moins, il faudrait ne pas utiliser la valeur maximale de c^R mais la valeur maximale de i^R, soit $I^R - 1$, pour trouver la valeur maximale de c^W.

Ce problème est supprimé si on utilise la *contrainte SIMD* telle qu'elle est définie dans la thèse de Nicolas Museux, page 44. Il est à revoir si on essaie de revenir à la contrainte dite SPMD.

Dans le cadre de DREAM-UP, il est envisageable de résoudre ce problème en augmentant systématiquement les dimensions des tableaux comme le fait THOMSON pour des transferts qui ne sont pas des multiples de quatre. Du fait des contraintes d'alignements et des formules de timing, cette approximation pourrait ne modifier pour THOMSON ni les temps d'exécution ni les capacités mémoire nécessaires (ATTENTION: 1-D pour THOMSON ou 2-D???).

7.4 Changement de repère

Deux blocs dépendent l'un de l'autre s'ils contiennent une paire d'itérations dépendant l'une de l'autre.

Si les calculs sont effectués sur $\mathbb{Q}$ ou $\mathbb{R}$, le passage du sommet $s_i = (r, w)$ au sommet $s_c = (c^R, c^W)$ se fait par une simple division.

Comme cela est montré en section 4.5.1, les composantes r de s_i sont données par les bornes de boucles I^R. Les composantes w_j de w sont données soit par le motif de lecture, soit par la combinaison de deux expressions d'accès au tableau X, qui sont toutes deux des contraintes supposées plus fortes que I^W:

$$\begin{array}{rcl} {\Omega_{pav}}_{k,j}^{W} w_j & = & K_k^R - M_k^W + 1 \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{fit}}_{k.k}^{R} (M_k^R - 1) + K_k^R \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{pav}}_{k,i}^{R} r_i + K_k^R - M_k^W + 1 \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{pav}}_{k,i}^{R} r_i + {\Omega_{fit}}_{k.k}^{R} (M_k^R - 1) + K_k^R \end{array}$$

Dans le premier cas, on a deux valeurs possible pour w_j. Elles sont données par les deux premières équations. Dans le deuxième cas, il faut substituer r_i par ses deux valeurs extrémales, 0 et $I^R - 1$:

$$\begin{array}{rcl} {\Omega_{pav}}_{k,j}^{W} w_j & = & K_k^R - M_k^W + 1 \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{fit}}_{k.k}^{R} (M_k^R - 1) + K_k^R \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{pav}}_{k,i}^{R} (I_i^R - 1) + K_k^R - M_k^W + 1 \\ {\Omega_{pav}}_{k,j}^{W} w_j & = & {\Omega_{pav}}_{k,i}^{R} (I_i^R - 1) + {\Omega_{fit}}_{k.k}^{R} (M_k^R - 1) + K_k^R \end{array}$$

Ces quatre équations définissent les sommets I, H, D et K du rapport DRET, de la thèse de Christophe Guettier[6] (p. 101) et de la thèse de Nicolas Museux. Le point A est soit confondu avec I, soit inexistant. Le point est soit confondu avec I, soit inexistant. Le point G est toujours confondu avec H ou inexistant par rapport au point H en l'absence de modulo. Le point F ne peut pas exister en l'absence de modulo sans qu'il y ait d'accès à des éléments non-définis. Le point E est soit confondu avec K, soit inexistant. Le point L et le point J sont en dehors du domaine de définition des itérations r. Le point C est soit confondu avec D soit inexistant.

En l'absence de motifs et/ou pour des partitionnements très particuliers, les points I et H d'une part, et les points D et K d'autre part, peuvent être confondus.

Pour des dimensions couplées (i,j,k), les points extrémaux dans les coordonnées initiales sont donc:

$$\begin{array}{rcl} (x_I, y_I) & = & \left(0, \dfrac{K_k^R - M_k^W + 1}{{\Omega_{pav}}_{k,j}^{W}}\right) \\ (x_H, y_H) & = & \left(0, \dfrac{{\Omega_{fit}}_{k.k}^{R} (M_k^R - 1) + K_k^R}{{\Omega_{pav}}_{k,j}^{W}}\right) \end{array}$$

6. Les notations utilisées pour les matrices d'accès ont été remplacées et rétuilisées pour dénoter les ensembles d'itérations.

$$
\begin{aligned}
(x_D, y_D) &= (I_i^R - 1, \frac{\Omega_{pav}{}_{k,i}^{R}(I_i^R - 1) + K_k^R - M_k^W + 1}{\Omega_{pav}{}_{k,j}^{W}}) \\
(x_K, y_K) &= (I_i^R - 1, \frac{\Omega_{pav}{}_{k,i}^{R}(I_i^R - 1) + \Omega_{fit}{}_{k.k}^{R}(M_k^R - 1) + K_k^R}{\Omega_{pav}{}_{k,j}^{W}})
\end{aligned}
$$

A FINIR. IL DOIT Y AVOIR PAS MAL DE BETISES QUANT AU NOMMAGE DES SOMMETS. MAIS IL N'Y A PAS D'UTILISATION DE C_MAX POUR AJOUTER DE L'IMPRECISION.

7.5 Conclusion sur la prise en compte du partitionnement

Faire le changement de coordonnées a priori n'est pas en soi une bonne idée puisqu'il conduit à ajouter potentiellement des itérations et des éléments de tableaux inexistant ou bien à ajouter des quantificateurs supplémentaires et/ou des contraintes supplémentaires.

Si on pouvait conserver une formulation rationnelle, la prise en compte du partitionnement pourrait se faire par un simple changement de repère sur les sommets rationnels du polytope de dépendance entre itérations élémentaires.

Malheureusement, les divisions introduites augmentent les artefacts entiers déjà dus aux motifs, les controlent par les variables du partitionnement et non par les constantes des motifs, et APOTRES ne peut fonctionner correctement qu'avec des sommets entiers.

8 Prise en compte des sommets entiers

Nicolas Museux a montré dans sa thèse qu'on ne pouvait pas utiliser des sommets rationnels pour calculer l'ordonnancement. Le solveur de contraintes utilisé, CLAIRE/ECLAIR, travaille sur des domaines finis comme les entiers positifs bornés et non sur des domaines infinis comme les rationnels, bornés ou non. Nicolas Museux a aussi montré qu'on ne pouvait pas remplacer un rationnel par une paire d'entiers (Section 2.2, Chapitre 5).

8.1 Exemple

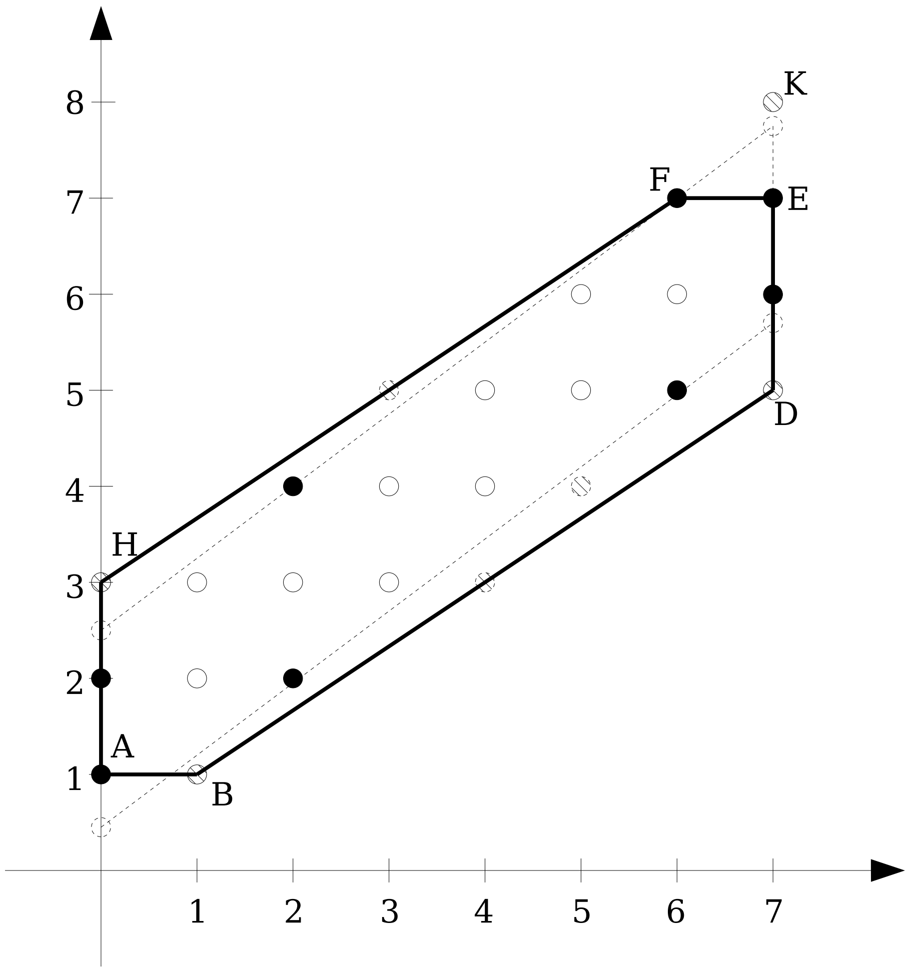
Il faudrait donc s'appuyer sur des sommets entiers. La figure 4 montre que le nombre de sommets entiers, qui sont représentés par des disques noirs, peut atteindre le double du nombre de sommets rationnels.

La figure 5 montre aussi que les nouveaux sommets entiers ne sont pas placés comme les points pris en compte dans l'approximation courante. Ces sommets sont représentés par des cercles pointillés et labellés par des lettres. Il manque un point au voisinage de H et un autre au voisinage de D. Le point B n'a pas forcément la même ordonnée que le point A. Il en est de même des points E et F.

On voit aussi sur cette figure que les sommets entiers utilisés par APOTRES définissent un polyèdre contenant toutes les dépendances, plus quelques dépendances fictives représentées par des disques hachurés. Cependant, ce polyèdre approximatif à sommets entiers ne contient pas le polyèdre rationnel des dépendances: deux des sommets de ce dernier se trouvent à l'extérieur.

Le nombre de sommets entiers est variable, mais il y en a huit dans le cas général et non six comme dans l'approximation APOTRES. Si les sommets rationnels sont aussi entiers, seuls quatre sommets sont utiles.

Fig. 5 – *Huit sommets entiers d'un parallélogramme rationnel*



8.2 Formulation symbolique des coordonnées des sommets entiers

On veut obtenir des formulations symboliques des coordonnées des sommets, telles que les sommets éventuellement excédentaire soient confondus avec les seuls sommets utiles.

Il s'agit de résoudre l'équation diophantienne $ay = bx + c$ et d'en trouver la plus petite solution positive. Suivant les valeurs et les modulos de a, b et c, il est parfois possible de trouver les solutions analytiques simples, mais le cas général impose le calcul des coefficients de Bezout.

Les coefficients a et b sont des fonctions du partitionnement. On ne peut donc en connaître ni les valeurs, ni les modulos des uns par rapports aux autres. Dans une approche manuelle, on utiliserait ce problème pour contraindre le partitionnement en demandant que a soit un multiple de b ou que b soit un multiple de a ou que a et b soit des multiples d'une taille de bloc élémentaire. Cette heuristique de partitionnement a été discutée plusieurs fois, mais elle n'a jamais été utilisée. La prendre en compte dans la résolution des contraintes d'ordonnancement dues aux dépendances n'est théoriquement pas acceptable. La prendre en compte dans le choix d'une approximation entière du polyèdre de dépendance pourrait être utile.

Nicolas Museux a réexploré ce problème dans la section 2 du chapitre 5 de sa these, a amélioré la sur-approximation choisie dans le rapport DRET, mais a conclu qu'il était illusoire de vouloir trouver une solution symbolique en présence de coefficients variables, comme c'est le cas quand le partitionnement est pris en compte.

8.3 Augmentation du nombre de sommets entiers

On pourrait envisager de calculer de nombreux points extrema du polytope de dépendance puisque les contraintes redondantes ne perturbent pas l'ensemble des solutions. On sait que les sommets entiers dont on a besoin se trouvent à proximité des sommets rationnels puisque les sommets entiers sont des solutions périodiques de période $(\Omega_{pav}^{R} L^{R} P^{R}, \Omega_{pav}^{W} L^{W} P^{W})$. Les matrices de partitionnement sont bornées par des constantes, les bornes de boucles I: $0 < LP \leq I$. De plus le nombre de points entiers dans le repère partitionné est borné par I et par I/LP. Ce nombre de points utiles x est donc inférieur à LP et à I/LP. Il est donc maximisable statiquement par la racine de I qui sature simultanément les deux contraintes.

En pratique, il faudrait prendre en compte R et W et choisir de travailler avec l'ensemble le plus petit de I^{W} et I^{R}.

Il faudrait aussi savoir ce qui se passe pour les valeurs des coordonnées ou des abscisses pour lesquelles aucune dépendance n'existe.

8.4 Construction de sommets entiers approximatifs

L'approche retenue par THALES est la construction de sommets entiers approximatifs. Ces sommets entiers définissent un ensemble de dépendance qui contient à coup sûr l'ensemble des points entiers du polytope de dépendance.

Théorèmes et preuves... sans doute à récupérer dans le rapport DRET 1996?

Pour simplifier la procédure de construction, on suppose à nouveau qu'une colonne de Ω_{pav}^{R} a au plus un élément non-nul. Ceci permet de partitionner les contraintes et de ne jamais avoir à traiter plus de deux variables simultanément.

Nous proposons de construire les sommets itérativement en utilisant en priorité les contraintes les plus fortes qui sont les contraintes générées par les accès aux tableau, puis en utilisant les contraintes sur les domaines d'itérations pour trouver les coordonnées des sommets qui manqueront encore.

On commence avec un sommet $s = (r, w)$ dont toutes les coordonnées sont inconnues.

Pour toutes les dimensions $k \in [1, n_X]$:

- s'il existe i_k et j_k tels que $\Omega_{pav}{}^{R}_{k,i_k} \neq 0$ et $\Omega_{pav}{}^{W}_{k,j_k} \neq 0$, on utilise la procédure bidimensionnelle définie dans la thèse de Nicolas Museux page 73, raffinée page 120, éventuellement modifiée en remplaçant les occurences de $L^W P^W C^W_m ax$ et de $L^R P^R C^R_m ax$ respectivement par I^W et I^R (voir section 7.3) et en fixant K^R à zéro, pour multiplier le nombre de sommets s par six en définissant les deux coordonnées s_{i_k} et $s_{n_R+j_k}$.
- sinon s'il existe j_k tels que $\Omega_{pav}{}^{W}_{k,j_k} \neq 0$, on multiplie le nombre de sommets s par deux en utilisant les deux valeurs:

$$s_{n_R+j_k} = \left\lfloor \frac{K^R_k - M^W_k + 1}{\Omega_{pav}{}^{W}_{k,j} L^W_{k,k} P^W_{k,k}} \right\rfloor$$

$$s_{n_R+j_k} = \left\lceil \frac{\Omega_{fit}{}^{R}_{k,k}(M^R_k - 1) + K^R_k}{\Omega_{pav}{}^{W}_{k,j} L^W_{k,k} P^W_{k,k}} \right\rceil$$

- sinon, on ignore la contrainte k parce qu'elle est trivialement satisfaite.

Pour toutes les dimensions j dans l'intervalle $[1, n_W]$, si s_{n_R+j} est encore non défini, on multiplie le nombre de sommets par deux en posant $s_{n_R+j} = 0$ ou bien $s_{n_R+j} = \left\lceil \frac{I^W_j}{L^W_{j,j} P^W_{j,j}} \right\rceil$.

Pour toutes les dimensions i dans l'intervalle $[1, n_R]$, si s_i est encore non défini, on multiplie le nombre de sommets par deux en posant $s_i = 0$ ou bien $s_i = \left\lceil \frac{I^R_i}{L^R_{i,i} P^R_{i,i}} \right\rceil$.

La correction de cette procédure s'appuie sur l'analyse des contraintes du polytope de dépendance faite en section 4.4.

8.5 Conclusion sur les sommets entiers

On ne peut pas calculer formellement les coordonnées des sommets entiers (Preuve?).

On n'a pas défini de meilleure approximation des projections 2-D du polytope entier des dépendances que celle qui est déjà connue.

On peut réutiliser les sommets approximatifs des projections 2-D définis dans le rapport DRET de 1996 et la thèse de Christophe Guettier, puis améliorés dans la thèse de Nicolas Museux, pour construire des sommets entiers sur-approximatifs valables.

9 Cas non pris en compte

Deux caractéristiques des spécifications APOTRES ne sont pas prises en compte: l'adressage des tableaux avec opérateur modulo et les dimensions infinies du temps-réel.

9.1 Prise en compte des opérateurs modulo dans les expressions d'accès

Les arguments des opérateurs modulos sont des constantes numériques.

Les opérateurs modulo ne peuvent apparaître que dans les utilisations de tableaux, et non dans leurs définitions.

Ils font apparaître dans le polyèdre de dépendance des sommets dont les projections sur l'espace des itérations de lecture ne sont pas des sommets de l'hypercube I^R.

Ils nécessitent l'ajout d'un quantificateur supplémentaire dans la définition des dépendances.

Ils ajoutent un terme dans l'équation de dépendance. Ce terme, s'il est mis du côté de l'expression des lectures, viole la contrainte de positivité de tous les coefficients. S'il est mis du côté de l'expression des écritures, il viole la contrainte de définition unique des éléments de tableau.

Le sujet est traité dans la thèse de Nicolas Museux, p. 75.

Comme les opérateurs modulo ne sont utilisés dans aucune application DREAM-UP, le sujet est laissé ouvert.

9.2 Prise en compte d'une dimension infinie au plus par nid de boucles

Les dimensions infinies font passer des polytopes de dépendance à des polyèdres de dépendance. Aux sommets il faut ajouter un rayon et poser une contrainte de dépendance légèrement différente. Elle ne doit pas faire intervenir le terme constant δ^R.

La contrainte de capacité mémoire bornée ajoutée à la contrainte d'ordonnancement conduit à un très fort couplage entre les coefficients d'ordonnancement de ces dimensions. Toutes les tâches doivent progresser à des vitesses multiples les unes des autres.

Aucune application DREAM-UP ne contient de boucle non-borné. Ce point peut donc être laissé ouvert.

10 Conclusion

Ce rapport met en évidence les problèmes posés par l'implantation actuelle du calcul des sommets du polytope de dépendance après partitionnement. Il propose un embryon de suite de tests de non-régression pour vérifier le bon fonctionnement des modèles d'ordonnancement et de dépendance.

Il reprend l'ensemble des fondements théoriques nécessaires au passage d'un ensemble de dépendances portant sur des itérations au sommets entiers d'un polytope incluant toutes les dépendances après partitionnement, ou presque. Quelques preuves se trouvent en principe dans le rapport DRET de 1996 ou dans la thèse de Nicolas Museux.

Ce rapport se heurte par ailleurs aux mêmes limites que les travaux précédents, et en particulier la thèse de Nicolas Museux, Chapitre 5, Section 2.

La meilleure approche pour le contrat DREAM-UP semble être de s'appuyer sur l'algorithme de calcul des sommets rationnels du polytope de dépendance présenté ici et de l'adapter pour prendre en compte les approximations nécessaires pour certaines dimensions en réutilisant la technique définie pour APOTRES afin de calculer symboliquement des sommets entiers en 2-D.

D'autres travaux devraient être entrepris pour réduire l'espace de recherche des ordonnancements. Par exemple, les constantes γ introduisent des non-évènements dans le temps logique des qu'elles atteignent comme seuil la date de la dernière itération des itérations précédentes. Si toutes les taches sont exécutées séquentiellement, cet ordre d'exécution n'est pas modifié par une augmentation du délai logique. De telles idées nécessite une formalisation mathématique avant qu'elles puissent être utilisées pour accélérer la résolution.