Méthodes Géométriques

pour

l'Analyse et la Transformation Symboliques

de Programmes

Philippe Clauss

ICPS
Université Louis Pasteur

Strasbourg

Méthodes géométriques pour l'analyse et la transformation symboliques de programmes

$\boxed{1}$ **Le projet de** ***système géométrique symbolique***

→ Schéma général

→ Types de problèmes traités

$\boxed{2}$ **Le moteur géométrique**

→ Sommets paramétriques
Domaines de validité

→ Polynômes d'Ehrhart

$\boxed{3}$ **La modélisation symbolique des problèmes**

→ Quelques modélisations opportunes

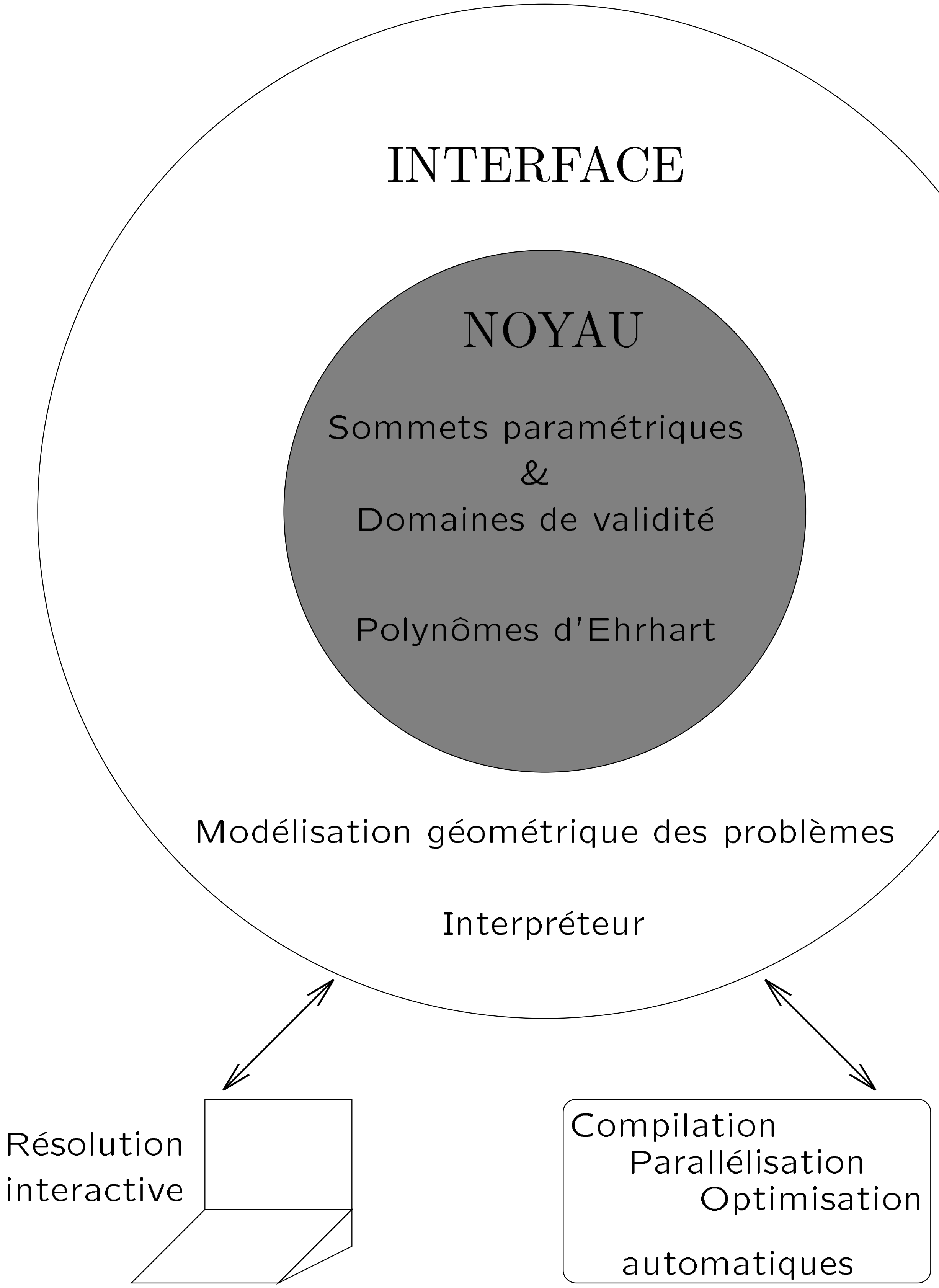
$\boxed{4}$ **Exemples d'application**

→ Analyse exacte des dépendances

→ Partitionnement paramétrique de boucles

→ Gestion des caches

Système Géométrique Symbolique



Problèmes traités :

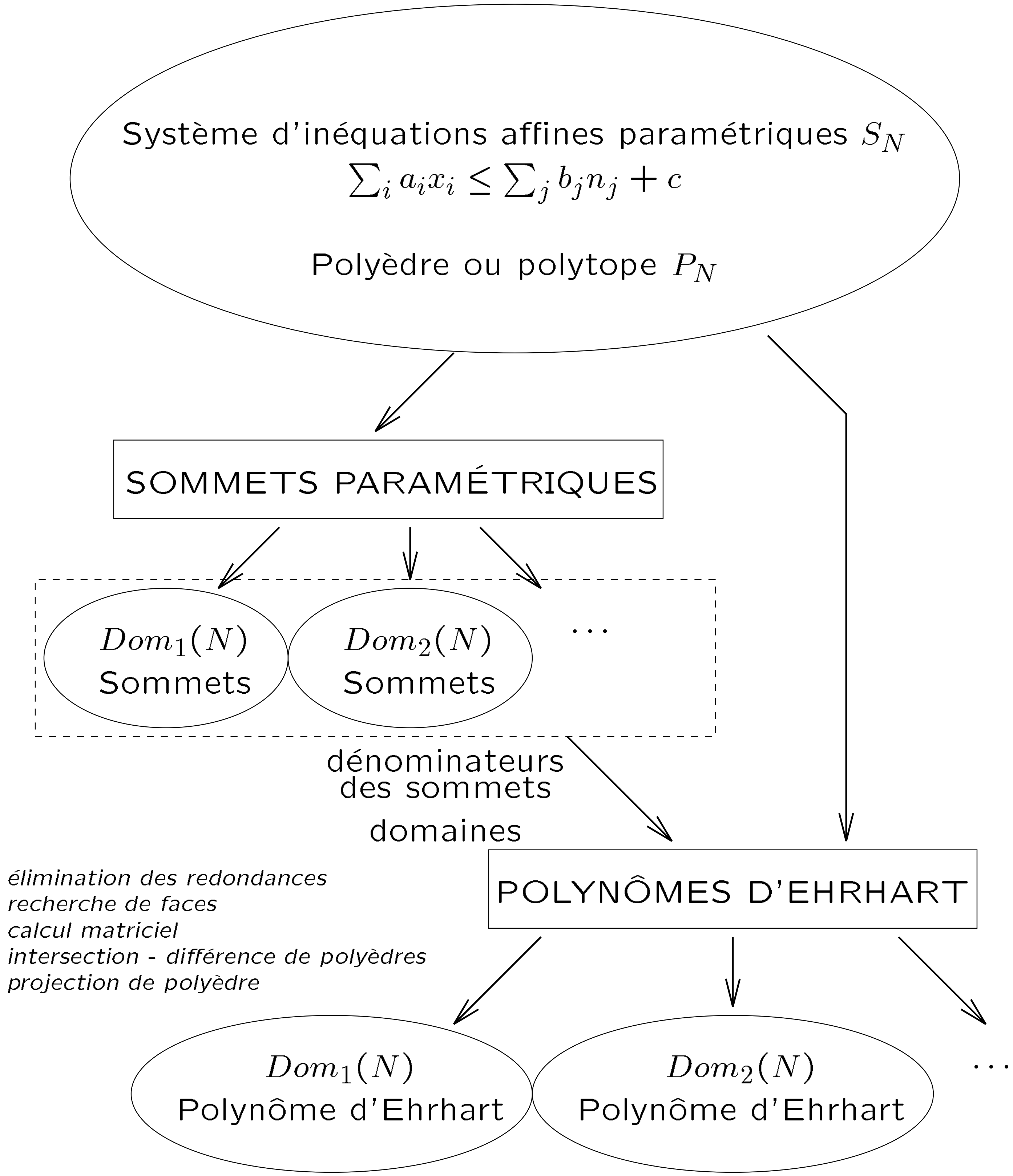
Analyse, tranformation et parallélisation de nids de boucles et de tableaux

- **Analyses symboliques :**
  - Analyse exacte des dépendances
  - Calcul du parallélisme maximum
  - Nombre d'éléments de tableau accédés
  - Nombre de lignes de cache allouées
  - Volume de communication
  - Nombre de processeurs virtuels alloués
  - Effets d'une distribution d'éléments de tableaux
  - ...

$\rightarrow$ **Transformations symboliques :**

$\rightarrow$ Boucles de parcours à bornes paramétrées
$\rightarrow$ Optimisation de la localité des données
$\rightarrow$ Minimisation des communications
$\rightarrow$ Ordonnancement temporel par sous-domaine d'indices de boucle
$\rightarrow$ Exploitation du parallélisme maximum
$\rightarrow$ Partitionnement, pavage optimal
$\rightarrow$ Optimisation des caches et lignes de cache
$\rightarrow$ Suppression de code inutile
$\rightarrow$ ...

Moteur géométrique



Moteur géométrique

Résolution de problèmes géométriques et arithmétiques généraux

$\rightarrow$ programmation linéaire paramétrique en nombres entiers (maximisation & minimisation)

$\rightarrow$ existence d'une solution entière d'un système d'équations ou d'inéquations paramétriques en nombres entiers

$\exists\, x \in \mathcal{Z}, f(x) = 0 \Leftrightarrow \{x \in \mathcal{Z} \mid -n \leq x \leq n, f(x) = 0\} \neq \emptyset$

$\rightarrow$ nombre symbolique de points entiers d'une transformation affine d'un polytope paramétré (projection, ...)

$\rightarrow$ analyse lexicographique

$\rightarrow$ véracité ou dénombrement des solutions d'un système d'équations ou d'inéquations paramétrique avec quantificateurs

$$\exists\, x \in D, Ax \leq B \Leftrightarrow D \cap \{x \in \mathcal{Z} \mid Ax \leq B\} \neq \emptyset$$
$$\forall\, x \in D, Ax \leq B \Leftrightarrow D \cap \{x \in \mathcal{Z} \mid Ax > B\} = \emptyset$$

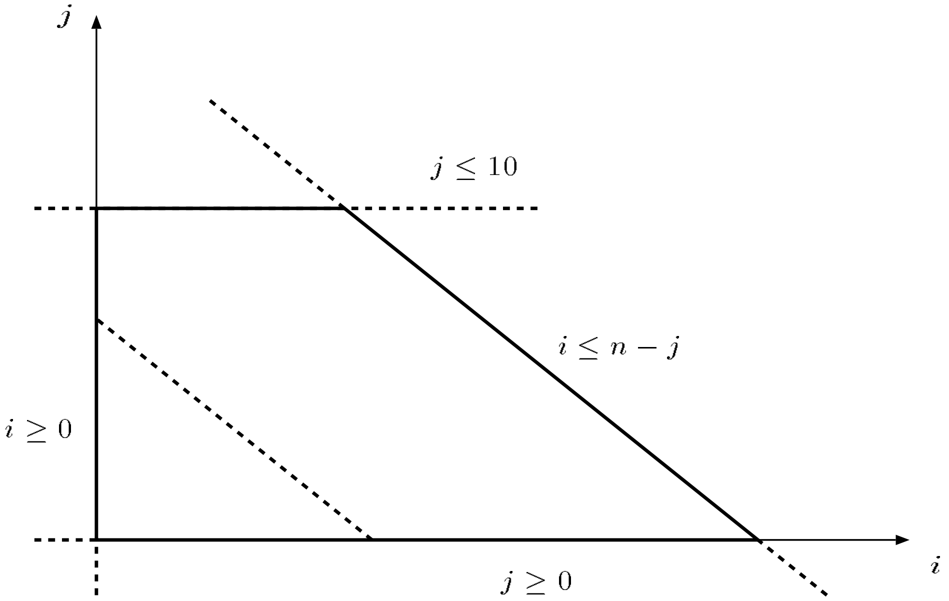
Polyèdre paramétré

Famille de polyèdres dépendant de $N \in \mathbb{Q}^m$:

$$P_N = \{x \in \mathbb{Q}^d \mid Ax \geq BN + C\}$$

Exemple : $P_n = \begin{pmatrix} i \\ j \end{pmatrix} \in \mathbb{Q}^2$

$$\begin{pmatrix} 1 & 0 \\ -1 & -1 \\ 0 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} i \\ j \end{pmatrix} \geq \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix} n + \begin{pmatrix} 0 \\ 0 \\ 0 \\ -10 \end{pmatrix}$$

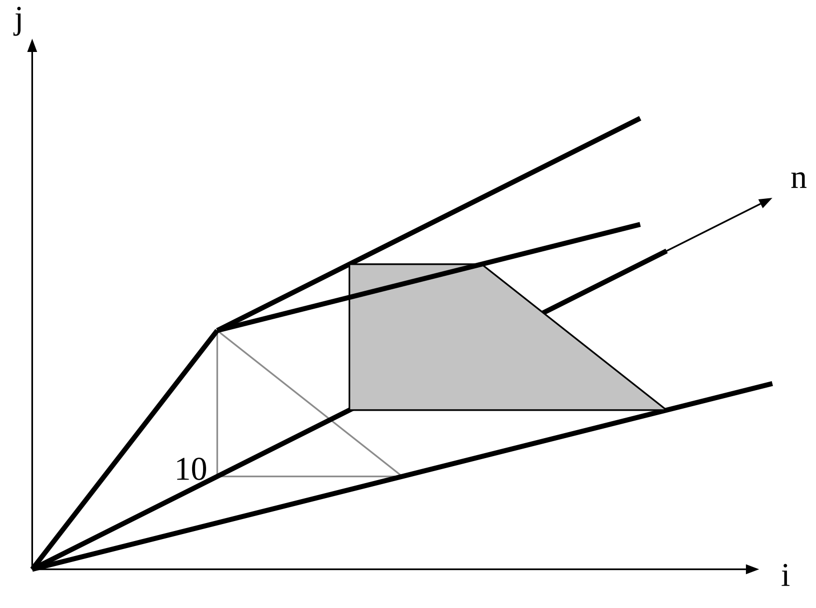


Polyèdre combiné

$$P' = \left\{ \begin{pmatrix} x \\ N \end{pmatrix} \in \mathbb{Q}^{p+m} \middle| [A \mid -B] \begin{pmatrix} x \\ N \end{pmatrix} \geq C \right\}$$

Exemple :

$$\begin{pmatrix} i \\ j \\ n \end{pmatrix} \in \mathbb{Q}^3 \quad \begin{pmatrix} 1 & 0 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} i \\ j \\ n \end{pmatrix} \geq \begin{pmatrix} 0 \\ 0 \\ 0 \\ -10 \end{pmatrix}$$



Généralisation

Théorême 1 :
Chaque sommet paramétrique de P_N correspond à une m-face unique du polyèdre combiné P' associé à P_N.

Théorême 2 :
Le domaine de validité sur N de chaque sommet paramétrique de P_N est la projection de la m-face de P' sur l'espace des paramètres.

Exemple :
- Si $0 \leq n \leq 10$ alors il y a 3 sommets:

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} n \\ 0 \end{pmatrix} \text{et} \begin{pmatrix} 0 \\ n \end{pmatrix}$$

- Si $n \geq 10$ alors il y a 4 sommets:

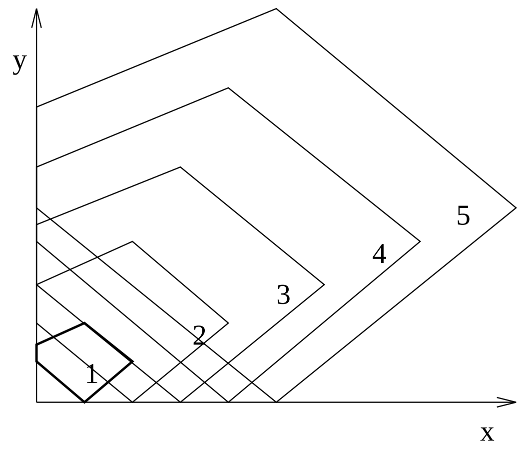
$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} n \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 10 \end{pmatrix} \text{et} \begin{pmatrix} n-10 \\ 10 \end{pmatrix}$$

Polyèdres homothétiques - bordés

= un polyèdre paramétré P_N défini sur un certain domaine de validité $Dom(N)$

Théorême 3 :
Le nombre symbolique de points entiers de P_N (sur $Dom(N)$) suit une relation de récurrence linéaire en N.

Exemple :



$$Dom(n) = \{n \geq 0\}$$

$$P_n \begin{cases} x - 2y \leq n \\ x + 2y \geq n \\ x \geq 0 \\ x - 4y \geq -3n \\ x + 2y \leq 3n \end{cases}$$

$$\begin{aligned} pts(n) = \; & pts(n-2) + 2pts(n-4) - 2pts(n-6) \\ & -pts(n-8) + 2pts(n-1) - 4pts(n-3) \\ & +4pts(n-7) - 2pts(n-9) + pts(n-10) \end{aligned}$$

Polynômes d'Ehrhart

Théorême 4, Polynôme d'Ehrhart :
Le nombre symbolique de points entiers d'un polyèdre homothétique-bordé P_N est un polynôme en N de degré d si P_N est entier (sommets entiers) et est un pseudo-polynôme en N de degré d si P_N est rationnel (au moins 1 sommet rationnel).

Exemple :

$$P_n \begin{cases} x - 2y \leq n \\ x + 2y \geq n \\ x \geq 0 \\ x - 4y \geq -3n \\ x + 2y \leq 3n \end{cases}$$

$$Dom(n) = \{n \geq 0\}$$

Sommets $(n, n), \left(0, \frac{3}{4}n\right), \left(0, \frac{n}{2}\right), \left(2n, \frac{n}{2}\right), (n, 0)$

$$EP(n) = \frac{9}{8}n^2 + n + \left[1, -\frac{1}{8}, \frac{1}{2}, -\frac{1}{8}\right]$$

La modélisation symbolique des problèmes

Quelques modélisations opportunes

- **Cadencement affine**

$$\forall\ X \in D_N, step(X) = \vec{T} \cdot X + t$$

$\Rightarrow$ Nombre de pas de calculs
= polynôme d'Ehrhart de $step(D_N)$
(transformation affine de polytope)

$\Rightarrow$ Parallélisme à t_0
= polynôme d'Ehrhart $EP(t_0, N)$ de
$P_{t_0} = \{X \in D_N \mid \vec{T} \cdot X + t = t_0\}$

$\Rightarrow$ Parallélisme maximum
= maximisation de $EP(t_0, N)$ en t_0

- **Boucle de parcours (paramétrée) de P**

$\rightarrow$ changement de base
$i = f_i(X), j = f_j(X), \ldots$

$\rightarrow$ calcul des sommets et domaines de validité de P_i
$\Rightarrow$ valeurs min et max de i par domaine

$\rightarrow$ calcul des sommets et domaines de validité de $P_{i,j}$
$\Rightarrow$ valeurs min et max de j en fonction de i par domaine

$\rightarrow \ldots$

$\rightarrow$ valeurs min et max du dernier indice $=$ sommets

Exemple d'application

Analyse exacte des dépendances

```
for i := 1 to n do
  for j := i to n do
    a(i+j) := a(i-j+n) + 1
```

Q : Étant donnée une itération (i_0, j_0), quelles sont les itérations (i, j) qui utilisent $a(i_0 + j_0)$ calculé en (i_0, j_0) ?

$\Rightarrow$ $\boxed{1}$ Quelle est la 1ère itération (i, j) qui met à jour $a(i_0 + j_0)$ après (i_0, j_0) ?

$\Rightarrow$ $\boxed{2}$ Quelles sont les itérations (i, j) qui utilisent $a(i_0 + j_0)$ avant sa remise à jour ?

Modélisation polyédrique des itérations $I = (i, j)$ postérieures à $I_0 = (i_0, j_0)$ selon l'ordre lexicographique :

- Domaine d'itération :
  $$D_n = \{I \in \mathcal{Z}^2 \mid 1 \leq i \leq n, i \leq j \leq n\}$$
- Itérations postérieures tq $i = i_0$ et $j > j_0$:
  $$P_{1,I_0,n} = \{(i,j) \in D_n \mid i = i_0, j > j_0, I_0 \in D_n\}$$
- Itérations postérieures tq $i > i_0$:
  $$P_{2,I_0,n} = \{(i,j) \in D_n \mid i > i_0, I_0 \in D_n\}$$

$\boxed{1}$ **Quelle est la 1ère itération (i, j) qui met à jour $a(i_0 + j_0)$ après (i_0, j_0) ?**

Modélisation polyédrique des itérations $I = (i, j)$ postérieures à $I_0 = (i_0, j_0)$ qui mettent à jour $a(i_0 + j_0)$:

$$\begin{aligned} P'_{1,I_0,n} &= \{(i,j) \in D_n \mid i = i_0, j > j_0, I_0 \in D_n, \\ &\qquad\qquad i + j = i_0 + j_0\} \\ &= \emptyset \\ P'_{2,I_0,n} &= \{(i,j) \in D_n \mid i > i_0, I_0 \in D_n, i + j = i_0 + j_0\} \end{aligned}$$

$\boxed{1}$

Recherche du minimum lexicographique dans $P'_{2,I_0,n}$

$\Rightarrow$ Recherche des sommets et domaines de validité de $P'_{2,i,I_0,n}$ (i en paramètre) :

$$\{1 \leq i_0 \leq j_0 - 2, j_0 \leq n, i_0 + 1 \leq i \leq \tfrac{i_0+j_0}{2}\}$$

$$1 \text{ sommet : } (-i + i_0 + j_0)$$

$\Rightarrow$ Minimum lexicographique :

$$(i_0 + 1, j_0 - 1)$$

$\Rightarrow$ **Conclusion :** si $i_0 \leq j_0 - 2$, $a(i_0 + j_0)$ est remis à jour en $(i_0 + 1, j_0 - 1)$, sinon, il n'est pas remis à jour

$\boxed{2}$ Quelles sont les itérations (i,j) qui utilisent $a(i_0+j_0)$ avant sa remise à jour ?

Modélisation polyédrique des itérations $I = (i,j)$ postérieures à $I_0 = (i_0, j_0)$ qui utilisent $a(i_0+j_0)$ avant sa remise à jour (si elle a lieu) :

- Mise à jour ou non :

$$P''_{1,I_0,n} = \{(i,j) \in D_n \mid i = i_0, j > j_0, I_0 \in D_n, \\ i - j + n = i_0 + j_0\}$$

- Mise à jour :

$$P''_{2.1,I_0,n} = \{(i,j) \in D_n \mid i > i_0, I_0 \in D_n, i_0 \leq j_0 - 2, \\ i = i_0 + 1, j < j_0 - 1, i - j + n = i_0 + j_0\}$$

- Pas de mise à jour :

$$P''_{2.2,I_0,n} = \{(i,j) \in D_n \mid i > i_0, I_0 \in D_n, \\ i_0 > j_0 - 2, i - j + n = i_0 + j_0\}$$

$\boxed{2}$

Recherche des sommets et domaines de validité de $P''_{1,I_0,n}$:

$$\{1 \leq i_0 \leq j_0 \leq \frac{n-1}{2}\}$$

$$\text{Sommet : } (i_0, n - j_0)$$

Recherche des sommets et domaines de validité de $P''_{2.1,I_0,n}$:

$$\{i_0 \geq 1, \frac{n+3}{2} \leq j_0 \leq n - i_0\}$$

$$\text{Sommet : } (i_0 + 1, n - j_0 + 1)$$

Recherche des sommets et domaines de validité de $P''_{2.2,I_0,n}$ (i en paramètre) :

$$\{1 \leq i_0 \leq j_0 \leq i_0 + 1 \leq i \leq i_0 + j_0 \leq n\}$$

$$\text{Sommet : } (i + n - i_0 - j_0)$$

$\Rightarrow$ **Conclusion :** L'élément $a(i_0 + j_0)$ calculé en (i_0, j_0) est utilisé par les itérations $I = (i, j)$ définies par :

Si $i_0 \leq j_0 - 2$

alors $I = (i, i + n - i_0 - j_0), i = i_0..i_0 + 1$

Sinon $I = (i, i + n - i_0 - j_0), i = i_0..i_0 + j_0$

Domaine d'utilisation de $a(i_0 + j_0)$ calculé en (i_0, j_0) :

$$Util_{i_0,j_0,n}$$

$$=$$

$$\{(i, i + n - i_0 - j_0) \in D_n \mid i_0 \leq i \leq i_0 + 1, i_0 \leq j_0 - 2\}$$
$$\cup$$
$$\{(i, i + n - i_0 - j_0) \in D_n \mid i_0 \leq i \leq i_0 + j_0, i_0 > j_0 - 2\}$$

$\Rightarrow$ **Volume de communication** maximum induit par une itération (i_0, j_0)

$=$ Polynôme d'Ehrhart de $Util_{i_0,j_0,n}$

$$EP(i_0, j_0, n) = \begin{cases} 2 & \text{si } i_0 \leq j_0 - 2 \\ j_0 + 1 & \text{si } i_0 > j_0 - 2 \end{cases}$$

$\Rightarrow$ Volume de communication total maximum

$= 2\times$ [nombre de points $(i_0, j_0) \in D_n \cap \{i_0 \leq j_0 - 2\}$]

$+ \sum_{j_0}$ [(nombre de valeurs de i_0 tq $(i_0, j_0) \in D_n \cap \{i_0 > j_0 - 2\}$) $\times (j_0 + 1)$]

$= 2\times$ [$EP(n)$ pour $D_n \cap \{i_0 \leq j_0 - 2\}$]

$+ \sum_{j_0}$ [($EP(j_0)$ pour $D_n \cap \{i_0 > j_0 - 2\}$) $\times (j_0 + 1)$]

$= n^2 - 3n + 2$

$+ \frac{5}{4}n^2 + \frac{11}{2}n + 2$

$= \frac{9}{4}n^2 + \frac{5}{2}n + 4$

Exemple d'application

Partitionnement paramétrique de boucle

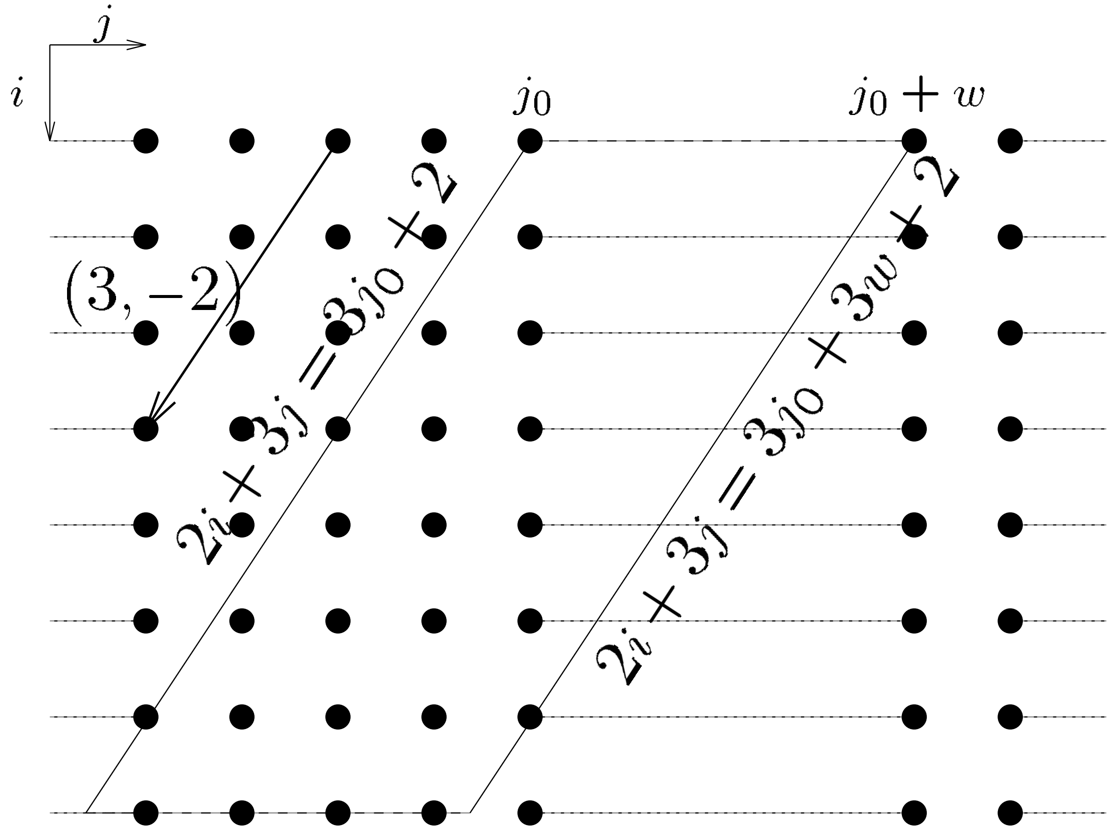
```
for i := 1 to 8 do
  for j := 1 to p do
    a(6i+9j-7) := a(6i+9j-7) + 5
```

→ Analyse de dépendance
$\Rightarrow (i+3, j-2)$ utilise $a(6i+9j-7)$
$(\vec{d} = (3, -2))$

→ Partitionnement en blocs non-communicants
$\Rightarrow$ allocation des itérations $(i+3, j-2)$ et (i, j) sur le même processeur

→ Partitionnement en parallélogrammes
- de hauteur 8 $(1 \leq i \leq 8)$
- de vecteurs directeurs $(3, -2)$ (non-comm.) et $(0, 1)$ (recouvrement)

$\Rightarrow$ Paramètre de largeur w



$\Rightarrow$ 1 parallélogramme = 1 boucle exécutée sur le processeur $(j_0 \div (w+1)) + 1$

$\Rightarrow$ Boucle de parcours : $x = 2i+3j, y = i+j$

for $x := 3j_0 + 2$ to $3j_0 + 3w + 2$ do

  for $y := \lceil \frac{x+1}{3} \rceil$ to $\lfloor \frac{x+8}{3} \rfloor$ do

    $a(3x-7) := a(3x-7) + 5$

$\rightarrow$ **Gestion de la mémoire cache**

$\Rightarrow$ pas de pb de **cohérence de cache** (1 processeur / élt a)

$\Rightarrow$ **localité temporelle** : (i,j) et $(i+3, j-2)$ sur le même processeur

Q : Étant donnée une taille de cache s, quelle est la taille w des parallélogrammes permettant de maximiser la localité des données accédées par la boucle (*cache miss*)?

$\Rightarrow$ Calcul du nombre de données accédées

$\rightarrow$ Domaine d'itération (parallélogramme)

$$D_{j_0,w} = \{(x,y) \in \mathcal{Z}^2 \mid 3j_0 + 2 \leq x \leq 3j_0 + 3w + 2, \frac{x+1}{3} \leq y \leq \frac{x+8}{3}\}$$

$\Rightarrow$ Polynôme d'Ehrhart de la transformation affine $T_{j_0,w}$ de $D_{j_0,w}$ par $z = 3x - 7$

$$T_{j_0,w} = \{z \in \mathcal{Z} \mid (x,y) \in D_{j_0,w}, z = 3x - 7\}$$

$\Rightarrow$ $EP(j_0, w) = 3w + 1$

Hyp : Éléments de tableaux codés sur 8 octets (double float)

$\Rightarrow$ $s = (3w + 1) \times 8$ octets

Hyp : $s = 1$ KB

$\Rightarrow$ $w = 42$

$\rightarrow$ **Localité spatiale**

$\Rightarrow$ allocation des données aux lignes de cache tq 2 données allouées à la même ligne de cache sont lues par 2 itérations successives
- (x, y) et $(x, y+1)$ lisent le même élément
- (x, y) et $(x+1, y')$ lisent resp. $a(3x-7)$ et $a(3x-4)$

Hyp : lignes de 32 octets (4 réels doubles)

$\Rightarrow$ fonction d'allocation
$a(k) \mapsto ligne[((k-7) \div 3 + 1) \div 4 + 1]$
$\Rightarrow$ Localité spatiale optimale

Hyp : $s = 1$ KB, $w = 42$
$\Rightarrow$ Nombre de lignes de cache allouées $= 32$ ($=$ minimum)

Conclusion

→ Objectifs équivalents à *Omega*

– approche arithmétique
  ≠ approche géométrique

– complexité *Omega* : $2^{2^{2^{\mathcal{O}(n)}}}$

– complexité *sommets paramétriques* :
  $\mathcal{O}\left(s^2 k^{\lfloor \frac{n+m}{2} \rfloor}\right)$

– complexité *polynômes d'Ehrhart* : $\mathcal{O}\left(v\left(dn\right)^m\right)$

s = nombre de sommets paramétriques
v = nombre de domaines de validité
d = dénominateur du polytope
n = nombre de variables
m = nombre de paramètres

→ Développements

– réalisés :
  * sommets paramétriques
  * polynômes d'Ehrhart

– futurs :
  * calculateur géométrique symbolique
  * interface de modélisation de problèmes types d'analyse de programmes