Factorization Forests of Finite Height

Jérémie Chalopin

26 Septembre 2002

Stage réalisé avec Hing LEUNG

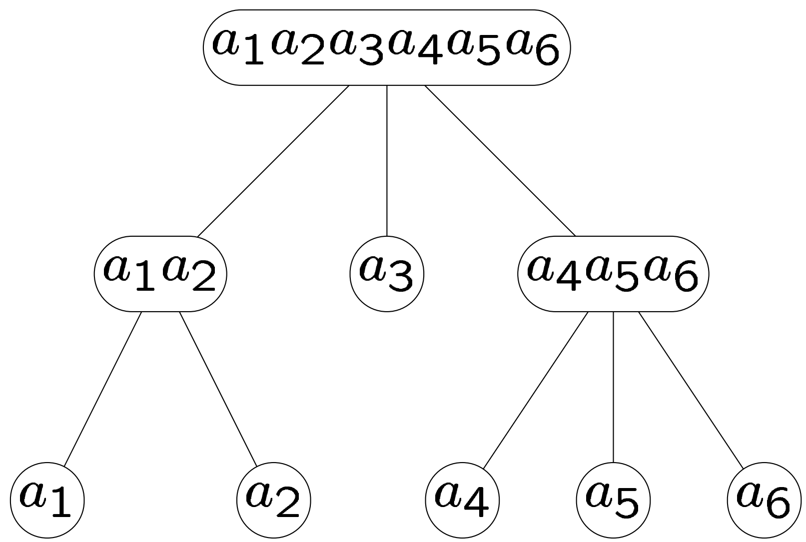
New Mexico State University, Las Cruces

Présentation du problème

Une fonction $d : A^+ \to (A^+)^+$ telle que :

- $d(x) = (x_1, x_2, \ldots, x_n) \Rightarrow x = x_1 x_2 \ldots x_n$
- $d(x) = x \Rightarrow x \in A$

Pourquoi parle-t-on de forêts ?



On définit le degré d'un sommet et la hauteur d'un arbre de manière usuelle.

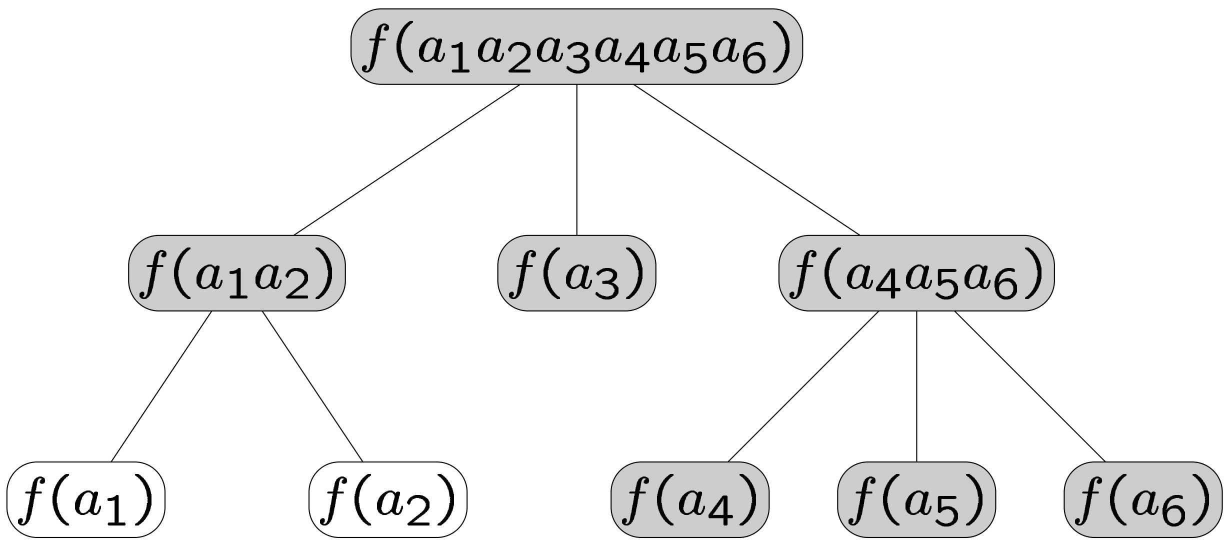
factorisation ramseyenne

Étant donné un morphisme $f : A^+ \to S$ où S est un semigroupe fini.

Pour tous mots $x, x_1, x_2, \ldots, x_n \in A^+$ tels que $d(x) = (x_1, x_2, \ldots, x_n)$

si $|d(x)| \geq 3$, il existe un idempotent $e \in S$ tel que

$$f(x) = f(x_1) = \ldots = f(x_n) = e$$

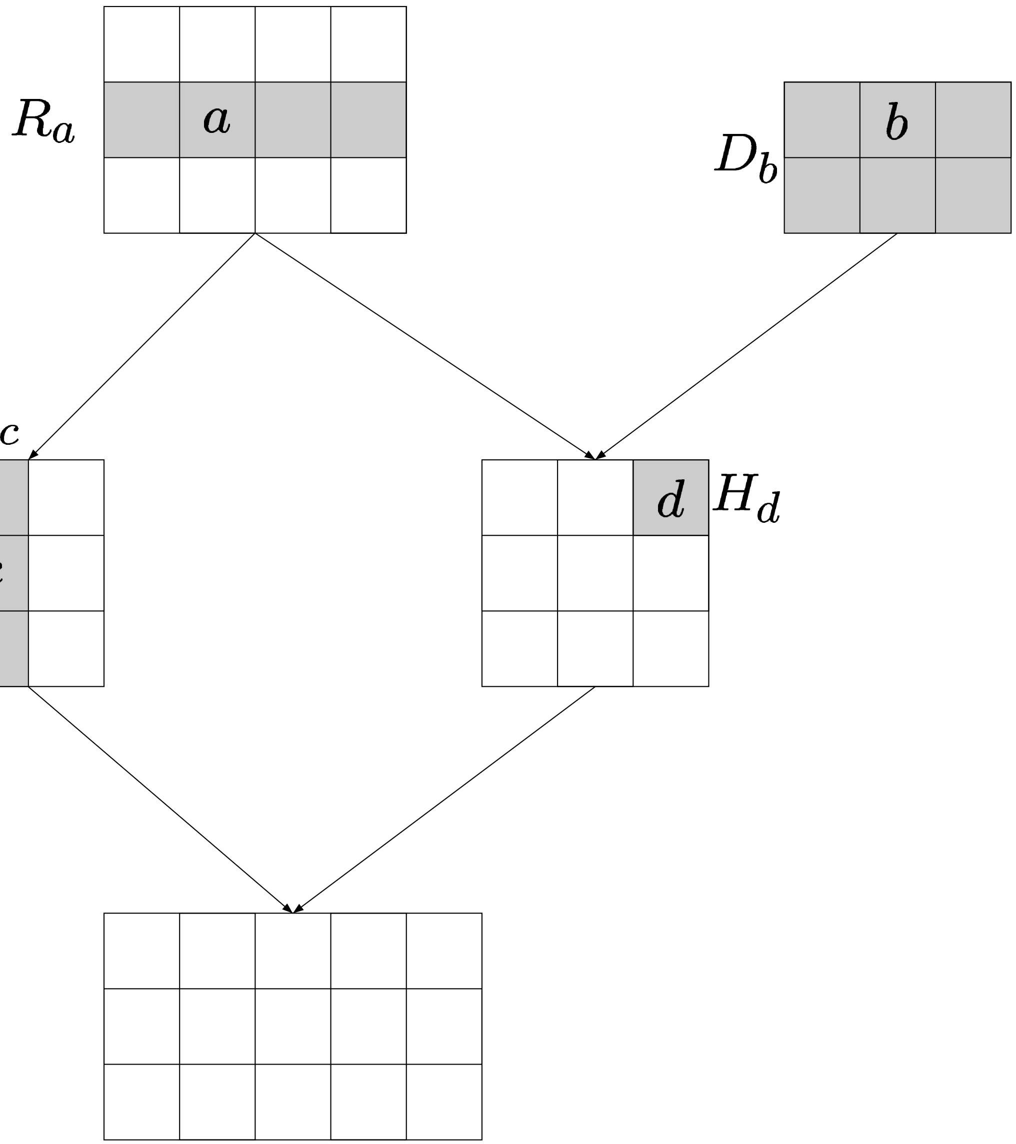

$f(a_1a_2a_3a_4a_5a_6)$
$f(a_1a_2)$
$f(a_3)$
$f(a_4a_5a_6)$
$f(a_1)$
$f(a_2)$
$f(a_4)$
$f(a_5)$
$f(a_6)$

Quelques résultats sur la théorie des semigroupes

Étant donné un semigroupe S, on obtient un monoide S^1

On définit les relations d'équivalences suivantes :

$$
\begin{aligned}
a\mathcal{R}b &\iff aS^1 = bS^1 \\
a\mathcal{L}b &\iff S^1a = S^1b \\
a\mathcal{D}b &\iff S^1aS^1 = S^1bS^1 \\
a\mathcal{H}b &\iff a\mathcal{R}b \wedge a\mathcal{L}b
\end{aligned}
$$


R_a
a
D_b
b
L_c
c
d
H_d

Lemme : Dans un semigroupe fini S, étant donnés $a, b \in S$, $a\mathcal{D}ab$ (resp. $b\mathcal{D}ab$) si et seulement si $a\mathcal{R}ab$ (resp. $b\mathcal{L}ab$).

Théorème : Dans un semigroupe fini S, soit D une $\mathcal{D}$-classe de S, H une $\mathcal{H}$-classe de D et $a, b \in H$, les points suivants sont équivalents

(i) $ab \in H$
(ii) H est un groupe
(iii) $ab \in D$

Théorème principal (Simon)

Théorème : Tout morphisme f d'un monoide libre A^+ dans un semigroupe fini S admet une forêt de factorisation ramseyenne de hauteur inférieure à $7|S|$.

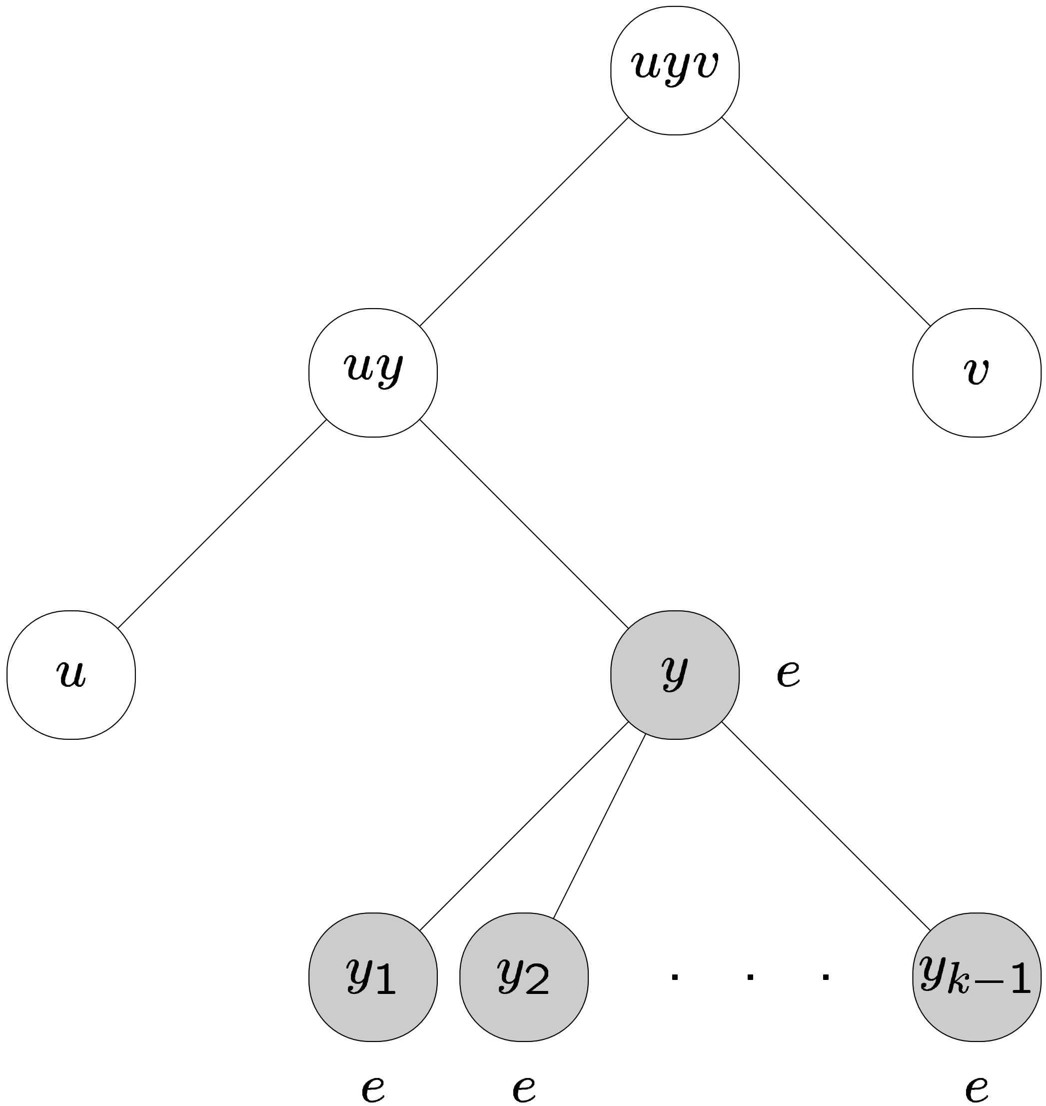
Le cas des groupes

Théorème : Tout morphisme f d'un monoide libre A^+ dans un groupe fini G admet une forêt de factorisation ramseyenne de hauteur inférieure à $3|G|$.

Étant donné un mot $x = a_1a_2\ldots a_p$, on note
$Pref(x) = \{f(u); u$ est un préfixe propre de $x\}$

Étant donné $b \in Pref(x)$, on recherche tous les préfixes s de x tels que $f(s) = b$.

$$x = \underbrace{\underbrace{\underbrace{\overbrace{a_1a_2\ldots a_{i_1}}^{u}}_{f(u)=b}\overbrace{a_{i_1+1}\ldots a_{i_2+1}}^{y_1}}_{f(y_1)=b}\ldots\overbrace{a_{i_{k-1}+1}\ldots a_{i_k}}^{y_{k-1}}}_{f(y_{k-1})=b}\overbrace{a_{i_k+1}\ldots a_p}^{v}$$


uyv
uy
v
u
y
e
y_1
y_2
$\cdot \quad \cdot \quad \cdot$
y_{k-1}
e
e
e

La complexité des mots des nouveaux noeuds a diminué :

- $b \notin Pref(u)$
- $\forall 1 \leq i \leq k-1, e \notin Pref(y_i)$
- $e \notin Pref(v)$

Par conséquent, on ne peut faire cette étape qu'au plus $|G|$ fois et on obtient donc un arbre de hauteur inférieure à 3$|G|$.

Dans une même $\mathcal{D}$-classe

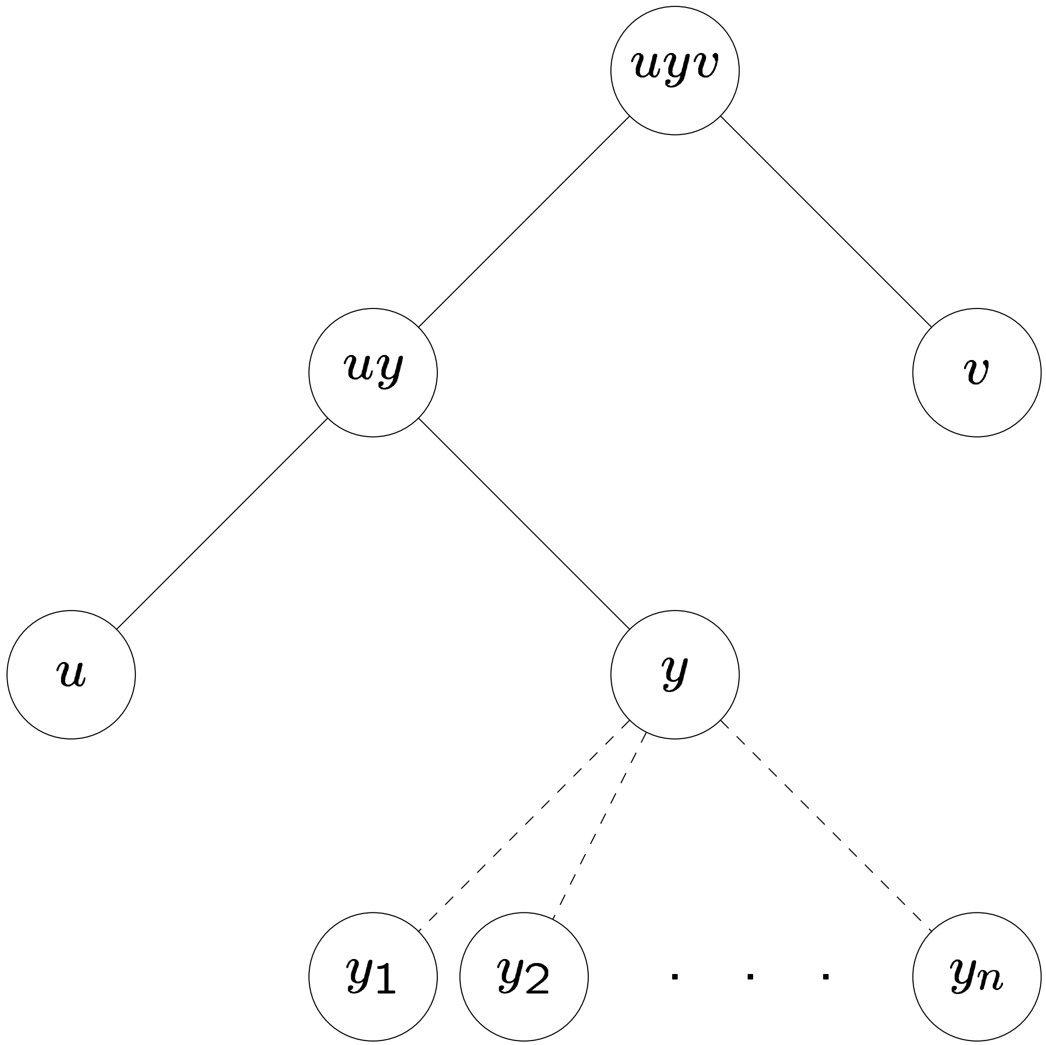
On va travailler sur des mots x tels que pour chaque lettre a de x, $f(a)$ et $f(x)$ sont dans la même $\mathcal{D}$-classe.

Théorème : Étant donné un mot x avec cette propriété, on peut trouver un arbre de factorisation ramseyenne de hauteur inférieure ou égale à $5|D_{f(x)}|$.

Étant donné un mot $x = a_1 a_2 \dots a_p$, on note $Int(x) = \{(L_{f(a_i)}, R_{f(a_{i+1})}); 1 \leq i \leq p-1\}$.

Soit $(l, r) \in Int(x)$, on décompose x comme suit :

$$x = \overbrace{a_1 \dots \underbrace{a_{i_1}}_{l}}^{u} \overbrace{\underbrace{\underbrace{a_{i_1+1}}_{r} \cdots \underbrace{a_{i_2}}_{l}}_{f(y_1) \in r \cap l}}^{y_1} \underbrace{a_{i_2+1}}_{r} \cdots \underbrace{a_{i_n}}_{l} \overbrace{\underbrace{\underbrace{a_{i_n+1}}_{r} \cdots \underbrace{a_{i_{n+1}}}_{l}}_{f(y_n) \in r \cap l}}^{y_n} \overbrace{\underbrace{a_{i_{n+1}+1}}_{r} \cdots a_p}^{v}$$


uyv
uy
v
u
y
y_1
y_2
$\cdot \quad \cdot \quad \cdot$
y_n

On note q la taille de chaque $\mathcal{H}$-classe.

A chaque étape, on obtient un arbre de taille $2 + 3q$.

On fait au plus une étape pour chaque couple (l, r) et donc on a un arbre plus petit que $5|D_{f(x)}|$.

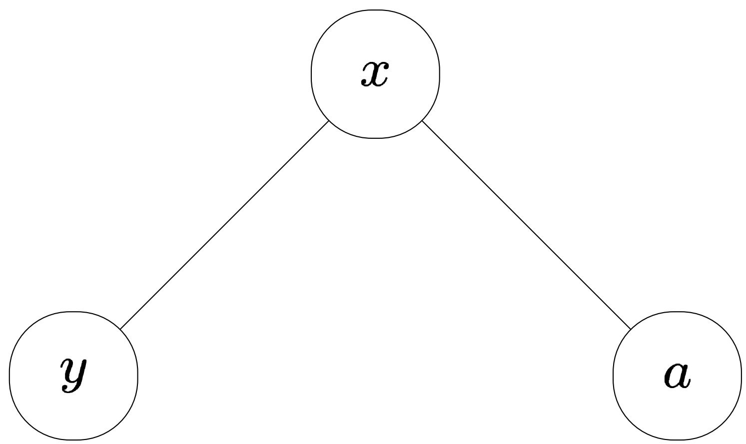
Cas général

Théorème : Pour tout morphisme $f : A^+ \to S$, d'un monoide libre dans un semigroupe fini, on peut trouver une forêt de factorisation ramseyenne de taille au plus $7|S|$.

Les $\mathcal{D}$-classes peuvent être partiellement ordonnées comme suit :

$$D_a \leq_{\mathcal{D}} D_b \iff S^1 a S^1 \subseteq S^1 b S^1$$

Si tous les préfixes propres u de x sont tels que $D_{f(u)} \neq D_{f(x)}$, soient $y \in A^+$ et $a \in A$ tels que $x = ya$.



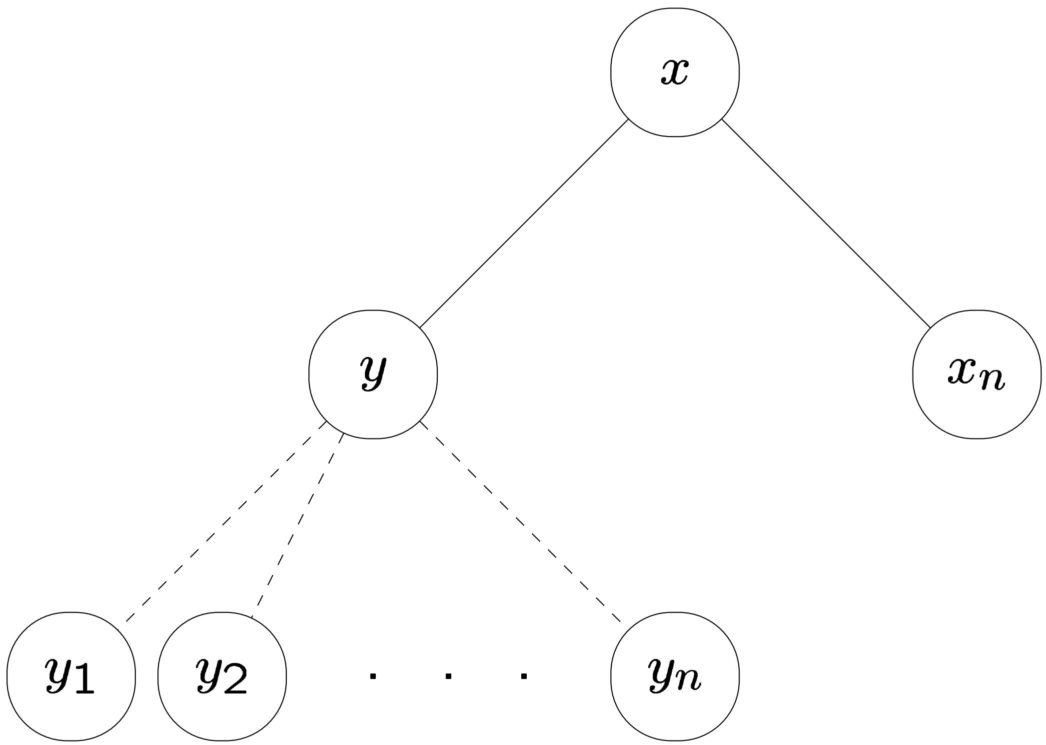
$$D_{f(x)} \leq_{\mathcal{D}} D_{f(y)}$$

Sinon, on décompose $x = a_1 \ldots a_p$ comme suit :

$$x = \underbrace{\overbrace{a_1 \ldots a_{i_1}}^{y_1}}_{f(y_1) \in D_{f(x)}} \ldots \underbrace{\overbrace{a_{i_n+1} \ldots a_{i_{n+1}}}^{y_n}}_{f(y_n) \in D_{f(x)}} \overbrace{a_{i_{n+1}+1} \ldots a_p}^{x_n}$$

Tous les y_i ont été traités dans le cas précédents

$D_{f(x)} \leq_{\mathcal{D}} D_{f(x_n)}$


x
y
x_n
y_1
y_2
$\cdot \quad \cdot \quad \cdot$
y_n

On fait au plus une fois cette étape pour chaque $\mathcal{D}$-classe.

Pour chaque $\mathcal{D}$-classe D, on crée un arbre de taille au plus $5|D|+2$, et donc on un arbre de factorisation ramseyenne de taille plus petite que $7|S|$.

Un exemple qui nécessite une taille au moins linéaire

On considère le semigroupe constitué des éléments $\alpha_1, \ldots \alpha_n$ avec la loi associative suivante : $\alpha_i \alpha_j = \alpha_j \alpha_i = \alpha_{max\{i,j\}}$.

On voit que chaque élément est dans sa propre $\mathcal{D}$-classe et que les $\mathcal{D}$-classes sont strictement ordonnées.

On considère l'alphabet $a_1 \ldots a_n$ et le morphisme f tel que $\forall 1 \leq i \leq n, f(a_i) = \alpha_i$.

On considère le mot $x_i = (((a_1{}^2 a_2)^2 a_3)^2 \ldots a_i)^2$ et on va montrer que la hauteur de tout arbre de factorisation ramseyenne pour x_i est de hauteur au moins i.

Pour $i = 1$, c'est évident.

Si $i \geq 2$, on ne peut pas avoir $|d(x)| \geq 3$

Donc $d(x) = (u, v)$ et x_{i-1} est un facteur de u ou de v et par récurrence la hauteur de l'arbre associé à u ou v est au moins $i - 1$ et donc la hauteur de l'arbre associé à x_i est au moins i.