PLUS COURTE DÉRIVATION

Idée Le but est de se donner, en fonction d'une grammaire et d'un texte / mot, une borne sur la longueur des dérivations à tester avant de conclure si oui ou non ce mot est engendré par cette grammaire.

Pté Soit $G = (\Sigma, \Gamma, R, S)$ une grammaire algébrique.
Soit $t = t_1 — t_{|t|}$ un mot de Σ^+.
Notons m le nombre maximal de symboles dans les termes droits des règles de G (s'il est supérieur à 2)

$t \in \mathcal{L}(G) \Leftrightarrow$ il existe une dérivation de S à t dans G de longueur $\leq m^{(1+|t|)\times|\Gamma|}$

Rappelons qu'une dérivation de S à t dans G correspond à un arbre de dérivation, dont la racine r est étiquetée par S et les feuilles par des lettres de Σ ou par ε de telle sorte que la concaténa° de leurs étiquettes de gauche à droite (la frontière de l'arbre) soit le mot t. Le nombre de dérivation correspond alors au nombre de noeuds internes.

Considérons l'arbre d'une plus courte dériva° dans G de S à t. Pour $i \in [1..n]$, une feuille f_i correspond à la lettre t_i, non seulem^t elle est étiquetée par la lettre t_i, mais c'est aussi la i-ème feuille étiquetée par une lettre en partant de la gauche.

Pour chaque sommet de l'arbre on peut regarder la frontière du sous arbre qu'il engendre. Cette frontière est un mot sur Σ, un facteur de t même. On affectera le sommet d'un attribut supplémentaire : son rang, défini comme la longueur de ce mot qu'il "couvre".

En particulier
- $\rightarrow rg(f_i) = 1$
- $\rightarrow rg(r) = |t|$ où r est la racine
- $\rightarrow rg(\pi(s)) \geq rg(s)$ où s est un sommet quelconque et $\pi(s)$ son père.

En considérant les sommets que l'on croise en remontant la branche de fi à r, le rang ne fait donc que croître et comme il passe de 1 à $|t|$ il ne peut croître strictem^t que $|t|-1$ fois.

On partage cette branche en blocs de sommets selon leur rang. On vient de dire qu'il y a au plus $|t|$ blocs

Hormis la feuille tous les sommets sont des noeuds internes étiquetés par des non terminaux, c-à-d des élém^t de Γ...

Chaque bloc a au plus $|\Gamma|$ noeuds (internes). En effet si un bloc a plus de $|\Gamma|$ noeuds, il en présente deux de m̂ étiquette, disons s_1 et plus bas s_2. Comme le rang ne change pas entre s_1 et s_2 les sommets intermédiaires ne produisent aucune lettre ; on pourrait donc recoller le sous arbre engendré par s_2 sous s_1, sans changer la frontière, en gardant un arbre de dérivation dans G. Cela nie la minimalité en nombre de noeuds / nbre de dérivations. IMPOSSIBLE

Donc une branche menant à une feuille étiquetée par une lettre a au plus $|\Gamma| \times |t|$ noeuds internes

Pour une branche menant à une feuille étiquetée par ε, on fait le m̂ raisonnem^t sauf que, puisque le rang de la feuille est nul, il y a au plus $|t|+1$ blocs et donc au plus $|\Gamma| \times (|t|+1)$ nds internes.

Cela borne la profondeur de l'arbre par $\underbrace{(|\Gamma| \times (|t|+1))}_{n} - 1$

Il y a 1 nd de prof. 0 (la racine)

Il y a au + m ———— 1 (ses fils)

———— m^2 ———— 2

⋮

———— m^{n-1} ———— $n-1$.

Au total il y a au plus $\sum_{i=0}^{n-1} m^i$ nds (internes)

or $\sum_{i=0}^{n-1} m^i = \frac{(1-m^n)}{(1-m)} = m^n \frac{\left(1/m^n - 1\right)}{(1-m)}$ car $m \geq 2$

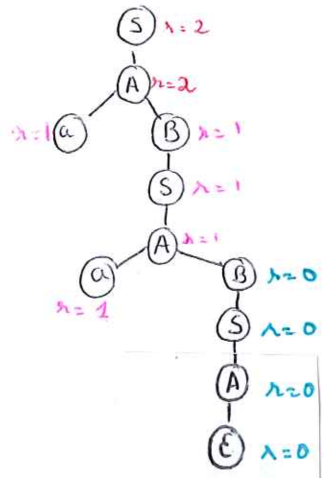
$$= \underbrace{\left(\frac{\overbrace{1 - 1/m^n}^{\leq 1}}{\underbrace{m-1}_{\geq 1}}\right)}_{\leq 1} m^n$$

D'où un nombre de dérivation minimal en $m^{(|t|+1) \times |\Gamma|}$

ex G : $S \to A$
$A \to aB \mid \varepsilon$
$B \to b \mid S$

$m = 2$
$|\Gamma| = 3$

$t = aa$ $|t| = 2$



Rq on voit bien que le $|t|+1$ est néc ici sur la branche qui va de S à ε. il y a $8 > |\Gamma| \times |t| = 6$ nds internes.