On remarque que dans l'expression rationnelle obtenue il apparaît plusieurs les même sous-expressions. Ainsi, si on représente les expressions rationnelles avec un DAG (directed acyclic graph, graphe acyclique), on aura une représentation plus compacte.

Borne supérieure

Proposition 7 *Si l'automate généralisé $\mathcal{A}$ a n états, alors l'expression régulière calculée est de longueur $O(4^n)$ mais son DAG contient $O(n^3)$ noeuds.*

Proof.

Soit u_k est le maximum des longueurs des expressions régulières à l'étape numéro k. Soit n_k le nombre de nœuds du DAG qui contient toutes les expressions régulières à l'étape numéro k.

On montre par récurrence sur k la propriété $\mathcal{P}(k)$ suivante :

$$u_k \leq 4^k(2(|\Sigma| + 1) + 4k)$$
$$n_k \leq 4k(n^2) + (n^3 + |\Sigma| + 1)$$

Intuitivement, dans la majoration de u_k, le 4^k correspond au fait que l'on recopie les expressions. Le $(2(|\Sigma| + 1))$ est une majoration des expressions rationnelles initiales et le $4k$ provient des 4 symboles (symbole $\cup$, symbole $*$ et deux symboles de concaténation) que l'on ajoute à chaque étape.

Intuitivement dans la majoration de n_k, le $(n^3 + |\Sigma| + 1)$ correspond au DAG de l'automate à l'étape 0. Puis, à chaque étape, on ajoute 4 nouveaux nœuds par transitions (et il y a n^2 transitions au plus).

- Cas de base. A l'étape 0, l'automate initial ne contient que des ϵ ou des a sur les transitions. Donc les expressions rationnelles dans l'automate généralisé sont de longueur au plus $2(|\Sigma| + 1)$. Donc $u_0 \leq 2(|\Sigma| + 1)$.

  Le DAG contient les nœuds ϵ et a pour toute lettre. Il contient aussi des noeuds sont les symboles $\cup$ des expressions rationnelles. Pour chaque transition, il y a au plus n symboles $\cup$. Il y a n^2 transitions. Il y a donc au plus $|\Sigma| + 1 + n^3$ nœuds.

  La propriété $\mathcal{P}(0)$ est donc vraie.

- Cas récursif. Considérons une étape $k > 0$. Supposons $\mathcal{P}(k-1)$ et montrons $\mathcal{P}(k)$.

  La nouvelle expression régulière sur une transition est de la forme

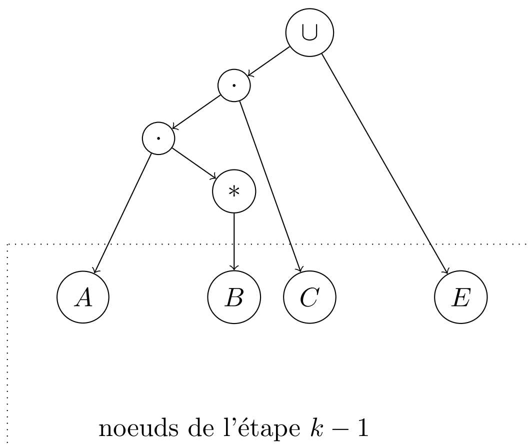
  $$((A.B^*).C) \cup E$$

  où A, B, C, E sont des expressions de l'automate de l'étape $k - 1$. On a $|((A.B^*).C) \cup E| = 1 + |E| + 1 + |C| + 1 + |A| + |B| + 1 = 4 + |A| + |B| + |C| + |E|$, le 4 correspond aux 4 symboles (une union, une étoile et deux concaténations. Donc $u_k \leq 4u_{k-1} + 4$.

On a donc par $\mathcal{P}(k-1)$,

$$\begin{aligned} u_k &\leq 4 \times 4^{k-1} \times (2(|\Sigma|+1) + 4(k-1)) + 4 \\ &\leq 4^k(2(|\Sigma|+1) + 4k) - 4^{k+1} + 4 \\ &\leq 4^k(2(|\Sigma|+1) + 4k). \end{aligned}$$

Le DAG à l'étape $k-1$ a n_{k-1} noeuds. Là, pour chaque transition, on ajoute 4 noeuds (pour les quatres symboles, étoile, l'union et deux concaténations) comme le dessin suivant :



Il y a n^2 transitions. Donc On a : $n_k = 4n^2 + n_{k-1}$.

Par $\mathcal{P}(k-1)$, on a $n_k \leq 4n^2 + 4(k-1)n^2 + (n^3 + |\Sigma| + 1) \leq 4kn^2 + (n^3 + |\Sigma| + 1)$.

D'où $\mathcal{P}(k)$.

Ainsi pour $k = n-2$, on voit que $u_n = O(4^n)$ et $n_k = O(n^3)$.
■