LEMME DE L'ÉTOILE

à finir

Idée : Dans un graphe à N sommets, un chemin de longueur N ($\begin{smallmatrix} N \text{ arêtes} \\ N+1 \text{ sommets}\end{smallmatrix}$) présente nécessairement une boucle.

I Conditions nécessaires

Cf. Sakarovitch I.1.3.c. p78-80

Soit L un langage sur l'alphabet Σ.

On introduit plusieurs propriétés d'existence de facteurs itérants :

$\bigstar_1$ Il existe $N \in \mathbb{N}^*$ tel que tout mot de longueur $\geq N$ de L se factorise en uvw où $\begin{cases} v \neq \varepsilon \\ u v^* w \subset L \end{cases}$

Rq on dit alors que v est un facteur itérant de $f = uvw$ pour L.

$\bigstar_2$ Il existe $N \in \mathbb{N}^*$ tel que tout facteur g de longueur $\geq N$ d'un mot fgh de L se factorise en uvw où
$\begin{cases} v \neq \varepsilon \\ f u v^* w h \subset L \end{cases}$

$\bigstar_3$ Il existe $N \in \mathbb{N}^*$ tel que pour tout mot $f \in L$ se décomposant en N blocs non vides $f = u v_1 v_2 \dots v_N w$ avec $v_i \neq \varepsilon$, il existe $\underset{j \leq k}{(j,k)} \in [1..N]$ tel que $u v_1 \dots v_{j-1} (v_j \dots v_k)^* v_{k+1} \dots v_N w \subset L$.

" Il y a une succession de blocs qui forme un facteur itérant.

lemme de l'étoile faible $\boxed{L \text{ reconnaissable} \Rightarrow \bigstar_1}$

lemme de l'étoile $\boxed{L \text{ reconnaissable} \Rightarrow \bigstar_2}$

lemme de l'étoile par bloc $\boxed{L \text{ reconnaissable} \Rightarrow \bigstar_3}$

En résumé $\boxed{L \text{ rec} \Rightarrow \bigstar_3 \Rightarrow \bigstar_2 \Rightarrow \bigstar_1}$

II Réciproques fausses

Cf. Sakarovitch p79 et p80 ex 1.24.

$\bigstar_1 \not\Rightarrow \bigstar_2$ $\quad U = \underbrace{\{a^n b^n \mid n \in \mathbb{N}\}}_{:= P_1} \cup \underbrace{A^* ba A^*}_{:= P_2}.$ $\quad A = \{a, b\}$.

$\boxed{U \text{ vérifie } \bigstar_1}$, en particulier pour $N = 1$.

En effet considérons $f \in U$ de longueur ≥ 1.

- Si $f \in P_1$, f s'écrit $a^M b^M$ où $M \geq 1$ (sinon $|f| = 0$)

La factorisation $\begin{cases} u = a^{M-1} \\ v = ab \\ w = b^{M-1} \end{cases}$ convient puisque $uw \in P_1 \subset U$, $uvw = f \in U$, pour $k \geq 2$, $uv^k w \subset P_2 \subset U$

donc autrement dit $\underline{uv^*w \subset U}$

- Sinon, $f \in P_2$ s'écrit $g ba h$ où $(g, h) \in (A^*)^2$.
  - Si $g \neq \varepsilon$ la factorisation $\begin{cases} u = \varepsilon \\ v = g \\ w = bah \end{cases}$ convient car $\forall n \in \mathbb{N}$, $v^n w = v^n ba h \in P_2 \subset U$
  - Si $h \neq \varepsilon$, de même $\begin{cases} u = gab \\ v = h \\ w = \varepsilon \end{cases}$ convient.
  - Sinon $f = ba$, et $\begin{cases} u = \varepsilon \\ v = ba \\ w = \varepsilon \end{cases}$ convient car $v^* \subset P_2 \subset U$.

$\boxed{U \text{ ne vérifie pas } \bigstar_2}$

En effet pour $N \in \mathbb{N}^*$, le mot $f = a^N b^N \in U$ car $\in P_1$

Pourtant son facteur $g = a^N$ qui est de longueur $\geq N$ n'admet que des factorisations de la forme $\begin{cases} u = a^\alpha \\ v = a^\beta \\ w = a^\gamma \end{cases}$ où $\beta > 0$

qui ne conviennent pas

car $uwb^N = a^{\alpha+\gamma} b^N \notin U$

en effet
- puisque $\alpha + \gamma < \alpha + \beta + \gamma = N$ $\quad uwb \notin P_1$
- puisque tous les b sont après les a, $uwb \notin P_2$

Quel que soit N, l'existence d'une bonne factorisation n'est pas assurée pour les facteurs.

$\star_2 \not\Rightarrow \star_3$

$V = \underbrace{\{(aab)^n(abb)^n \mid n \in \mathbb{N}\}}_{:= Q_1} \cup \underbrace{A^*\{aaa, bbb, aabb\}A^*}_{:= Q_2}$

où $A = \{a, b\}$.

• $\boxed{V \text{ vérifie } \star_2}$ en particulier pour $N = 6$.

Considérons $f \in V$ qui présente un facteur g tq $|g| \geq 6$.
En fait on va m̂ supposer que $|g| = 6$ quitte à enlever des lettres à g (un facteur itérant de $\tilde{g}$ ainsi obtenu sera bien un facteur du g initial)

- Si $f \in Q_1$ alors $f = f_1 g f_2$ où :

$g = aabaab$ $\quad u = aa, v = b, w = aab$
ou $g = abaaba$ $\quad u = abaa, v = b, w = a$
ou $g = baabab$ $\quad u = baa, v = b, w = aa$
ou $g = aababb$ $\quad u = aa, v = b, w = abb$

alors $f_1 u w f_2 \in A^* aaa A^* \subset Q_2 \subset V$
et $\forall k \geq 2$ $f_1 u v^k w f_2 \in A^* aabb A^* \subset Q_2 \subset V$
or pour $k = 1$ $f_1 u v w f_2 = f \in V$.

ou $g = ababba$ $\quad u = ab, v = a, w = bba$
ou $g = babbab$ $\quad u = b, v = a, w = bbab$
ou $g = abbabb$ $\quad u = abb, v = a, w = bb$
ou $g = bbabba$ $\quad u = bb, v = a, w = bba$

alors $f_1 u w f_2 \in A^* bbb A^* \subset V_2 \subset V$
et $\forall k \geq 2$ $f_1 u v^k w f_2 \in A^* aabb A^* \subset V$
or pr $k = 1$ $f_1 u v w f_2 = f \in V$

- Si $f \in Q_2$, f s'écrit $f_1 f_2 f_3$ avec f_1, f_3 quelconques
$f_2 = aaa$ ou bbb ou $aabb$.

↳ Si g ne chevauche pas f_2, il est facteur itérant sans pb.

↳ Si g chevauche f_2, on choisit v une partie de g hors de f_2, et c'est itérant sans pb.
Ça existe bien car $|g| = 6 > 4 = \max\{|aaa|, |bbb|, |aabb|\}$

$\boxed{V \text{ ne vérifie pas } \bigstar_3}$ Soit $N \in \mathbb{N}^*$.

On considère $f = (aab)^N (abb)^N \in Q_1 \subset V$, découpé par blocs de 3

$f = \underbrace{aab}_{v_1}, \underbrace{aab}_{v_2}, \dots \underbrace{aab}_{v_N}, \Big| \underbrace{abb}_{v_{N+1}}, \text{———} \underbrace{abb}_{v_{2N}}.$

Soit $g = v_i — v_j$ avec $i \leq N$, $j \geq N$. ~~[illegible]~~

g^2 fait apparaître du $v_j v_i = abbaab$, ce qui est exclu ds les mots de Q_1. Et comme $v_1 — v_{i-1} g^2 v_{j+1} — v_{2N}$ ne fait pas non plus apparaître aaa, ou bbb ou $aabb$ ($bbaa$ apparaît) donc $v_1 — v_{i-1} g^2 v_{j+1} — v_{2N} \notin Q_1 \cup Q_2 = V$, donc g n'est pas facteur itérant.

Soit $g = v_i — v_j$ avec $i \leq j \leq N$

$v_1 — v_{i-1} v_{j+1} — v_{2N}$ n'est pas dans Q_1 car il n'a pas autant de facteurs blocs aab et abb. Donc g n'est pas facteur itérant

De même si $g = v_i — v_j$ avec $N+1 \leq i \leq j$.

On peut donc construire un mot décomposé en bloc qui n'admet pas de factorisa par bloc avec fact. itérant, ayant un nombre de bloc arbitrairement grand. Donc $\bigstar_3$ n'est pas vérifiée.

$\mathcal{A}_3 \not\Rightarrow$ rec. $\quad A = \{a,b,c\}$ $\quad B = \{a,b,c,\#\}$.

$K = \{uu \mid u \in A^+\}$ $\quad K' = \{u\#v \mid (u,v) \in A^* \text{ avec } u \neq v\}$

$$L = K' \cup \underbrace{A^*KA^*\#A^*}_{:= L_1} \cup \underbrace{A^*\#A^*KA^*}_{:= L_2}$$

! $\boxed{L \text{ vérifie } \mathcal{A}_3}$! en particulier pour $N = 3$.

Notons déjà que tous les mots de L ont exactem^t 1 $\#$.
Des trois blocs considérés dans un mot de L, au moins deux seront sans $\#$. Dans la suite on ne considère que ceux là car un bloc contenant $\#$ ne peut être itérant du fait que chaque mot de L a exactem^t un $\#$.

On a alors trois cas possibles

⚠ ici la notat° bloc a bloc b est maladroite car a et b n'ont ici rien à voir avec les lettres de l'alphabet

- $f = \alpha \llcorner \text{bloc } a \lrcorner \llcorner \text{bloc } b \lrcorner \beta \# \gamma$

  - Si $\alpha \llcorner a \lrcorner \beta \# \gamma$ alors $\llcorner b \lrcorner$ est un bloc itérant.

    En effet $\alpha \llcorner a \lrcorner \beta \# \gamma \in K'$ et $\forall k \geq 2$ $\alpha \llcorner a \lrcorner (\llcorner b \lrcorner)^k \beta \# \gamma \in L_1$

  - Si $\alpha \llcorner b \lrcorner \beta \neq \gamma$ alors de m̂ $\llcorner a \lrcorner$ est bloc itérant

  - Sinon, comme $\alpha \llcorner a \lrcorner \beta = \gamma = \alpha \llcorner b \lrcorner \beta$, on en déduit $\llcorner a \lrcorner = \llcorner b \lrcorner$.
    et $\alpha\beta \neq \gamma$ donc $\llcorner a \lrcorner \llcorner b \lrcorner$ est facteur itérant.
    En effet $\alpha\beta\#\gamma \in K'$ et $\forall k \geq 2$ $\quad \alpha(\llcorner a \lrcorner \llcorner b \lrcorner)^k \beta \# \gamma \in L_1$

- $f = \alpha \# \beta \llcorner \text{bloc } a \lrcorner \llcorner \text{bloc } b \lrcorner \gamma$ $\quad$ Se traite symétriquem^t avec L_2 au lieu de L_1.

- $f = \alpha \llcorner \text{bloc } a \lrcorner \beta \# \gamma \llcorner \text{bloc } b \lrcorner \delta$

  - Si $\alpha \llcorner a \lrcorner \beta \neq \gamma\delta$ alors $\llcorner b \lrcorner$ est facteur itérant

    En effet $\alpha \llcorner a \lrcorner \beta \# \gamma\delta \in K'$ et $\forall k \geq 2$ $\quad \alpha \llcorner a \lrcorner \beta \# \gamma (\llcorner b \lrcorner)^k \delta \in L_2$

  - Si $\alpha\beta \neq \gamma \llcorner b \lrcorner \delta$ alors $\llcorner a \lrcorner$ est facteur itérant

    En effet $\alpha\beta\#\gamma \llcorner b \lrcorner \delta \in K'$ et $\forall k \geq 2$ $\quad \alpha(\llcorner a \lrcorner)^k \beta \# \gamma b \delta \in L_1$

  - le cas $\alpha \llcorner a \lrcorner \beta = \gamma\delta$ et $\alpha\beta = \gamma \llcorner b \lrcorner \delta$ est impossible car on aurait simultanement $|\alpha\beta| < |\gamma\delta|$ et $|\alpha\beta| > |\gamma\delta|$!

$\boxed{\text{L n'est pas reconnaissable}}$

On s'avance un peu, et on n'utilise ici pour ① que L n'est pas rec. que L^c ne vérifie pas $\bigstar_3$.

Soit $N \in \mathbb{N}$ — En fait on sait déjà que si L était rec, L^c le serait aussi et alors L^c vérifierait $\bigstar_3$.

~~On considère le mot $u_1 = abc$ et le mot $u_2 = bac$~~
~~et le mot $f = v \# v$ où $v = u_1 a\, u_1 u_2 a\, u_1 u_2 u_2 a \ldots u_1 (u_2)^N a$~~
~~Il est clair que $f \notin K'$.~~
~~De plus $\notin A^* K A^*$~~

On considère un mot u sans facteur carré (ie $u \notin A^*KA^*$) de longueur $\geqslant N$, et on le partage arbitrairem^t en N blocs.
On pose alors $f = u \# u$. $f \notin K'$ et puisque $\notin A^*KA$, $f \notin L_1$ et $f \notin L_2$. Donc $f \in L^c$.
Pourtant quels que soient les blocs choisis, qui permettent d'écrire $u = \alpha \underset{\substack{\uparrow \\ \text{blocs} \\ \text{choisi}}}{b} \gamma$, $\alpha\gamma \# u \in L$ autrement dit $\alpha b^0 \gamma \# u \notin L^c$.

Donc L^c ne vérifie pas $\bigstar_3$.

Rq On peut construire un mot sans facteur carré arbitrairem^t grand parce qu'on a 3 lettres.
Le procédé est le suivant.

$$\varphi = \begin{pmatrix} a \to abc \\ b \to ac \\ c \to b \end{pmatrix} \qquad u_0 = a, \quad u_{n+1} = \varphi(u_n) = \varphi(u_n^1)\varphi(u_n^2)\ldots\varphi(u_n^{|u_n|})$$

ex $u_0 = a$, $u_1 = abc$, $u_2 = abc\ ac\ b$

$u_3 = abc\ ac\ b\ abc\ b\ ac$

$u_4 = abc\ acb\ abc\ b\ ac\ abc\ ac\ b\ ac\ abc\ b$