Exemples de grammaires hors-contexte

Exemple 1 *Soit* $L = \{w \in \{a,b\}^* : |w| \equiv 0 \pmod 2\}$ *et soit* $G = \{\{S\}, \{0,1\}, R, S\}$ *une grammaire hors-contexte avec* $R = \{S \longrightarrow \varepsilon | S00 | S01 | S10 | S11\}$*. Alors* $L = L(G)$*.*

Démonstration. *On montre d'abord, par récurrence sur la longeur de* w*, que si* $w \in L$ *alors* $S \overset{*}{\Longrightarrow} w$*. Si* $|w| = 0$*, alors* $w = \varepsilon$ *et* $S \overset{*}{\Longrightarrow} \varepsilon$ *car* $S \longrightarrow \varepsilon$ *est une règle. Si* $|w| = k > 0$*, alors* $w = uxy$ *avec* $u \in \Sigma^*$*,* $|u| \equiv 0 \pmod 2$ *et* $x, y \in \Sigma$*. Par l'hypothèse de récurrence,* $S \overset{*}{\Longrightarrow} u$*. Donc* $S \overset{1}{\Longrightarrow} Sxy \overset{*}{\Longrightarrow} uxy$ *car* $S \longrightarrow xy$ *est une règle.*
Ensuite, on prouve, par récurrence sur la longueur k *de la dérivation que si* $w \in L(G)$ *alors soit* $w \in L$*, soit il existe un* $u \in L$ *tel que* $w = Su$*. Si* $k = 0$*,* $S \overset{0}{\Longrightarrow} S$ *et* $S = S\varepsilon$*, avec* $\varepsilon \in L$*. Si* $k = 1$*,* $S \overset{1}{\Longrightarrow} w$ *directement et alors* $w = \{\varepsilon, S00, S01, S10, S11\}$ *et est de la bonne forme. Si* $k > 1$*,* $S \overset{k-1}{\Longrightarrow} vSz \overset{1}{\Longrightarrow} w$*. Par l'hypthèse de récurrence,* $vSz = Su$*,* $u \in L$*. Donc* $v = \varepsilon$ *et soit* $w = u$ *en utilisant* $S \longrightarrow \varepsilon$*, soit* $w = Sxyu$ *en applicant* $S \longrightarrow Sxy$*,* $x, y \in \Sigma$*. Donc* $S \overset{*}{\Longrightarrow} w$ *avec* $w \in \Sigma^*$ *uniquement si* $w \in L$*.* □

Exemple 2 *Soit* $L = \{a^n b^n : n \in \mathbb{N}\}$ *un langage non-régulier sur* $\Sigma = \{a, b, c, d\}$*. Définissons la grammaire hors-contexte* $G = (\{S\}, \Sigma, R, S$ *avec* $R = \{S \longrightarrow \varepsilon | aSb\}$*. On prétend que* $L = L(G)$*.*
On répète la procédure utilisée dans le premier exemple. Pour commencer, on montre par récurrence sur $|w| = k$ *que* $L \subseteq L(G)$*. Notons que tout mot de* L *est de longueur paire. Si* $k = 0$*,* $w = \varepsilon \in L$ *et* $S \overset{1}{\Longrightarrow} \varepsilon$ *car* $S \longrightarrow \varepsilon$ *est une règle. Pour* $k > 1$*, si* $|w| = k$ *alors* $w = aub$ *for* $u \in L$*. Par l'hypthèse de récurrence,* $S \overset{*}{\Longrightarrow} u$*. Donc* $S \overset{1}{\Longrightarrow} aSb \overset{*}{\Longrightarrow} aub = w$ *et* $w \in L(G)$*.*
Pour montrer que $L(G) \subseteq L$*, on fait la récurrence sur* k*, le longueur de la dérivation. On montre que* $S \overset{k}{\Longrightarrow} w$ *si, et seulement si,* $w = a^{k-1} X b^{k-1}$ *avec* $X \in \{S, \varepsilon\}$*. Si* $k = 1$*,* $S \overset{*}{\Longrightarrow} \varepsilon$ *est de la bonne forme, et* $\varepsilon \in L$*. Pour* $k > 1$*,* $S \overset{k}{\Longrightarrow} w$ *si, et seulement si,* $S \overset{1}{\Longrightarrow} aSb \overset{k-1}{\Longrightarrow} aub = w$ *avec* $S \overset{k-1}{\Longrightarrow} u$*. Par l'hypothèse de récurrence,* $u = a^{k-2} X b^{k-2}$ *et on a alors* $S \overset{1}{\Longrightarrow} a^{k-1} X b^{k-1}$*. Par conséquent, si* $w \in L(G)$*,* $w \in L$ *et* $L(G) \subseteq L$*.*

Exemples de langages qui ne sont pas hors-contexte

Exemple 3 *(exemple canonique) Soit* $\Sigma = \{a, b, c\}$*. Le langage* $L_1 = \{a^n b^n c^n \in \Sigma^* : n \in \mathbb{N}\}$ *n'est pas hors contexte. Pour le voir, prenons* p*, la constante du lemme de pompage, et le mot* $a^p b^p$*, qui est bien évidemment dans* L*. Supposons que* L *soit hors-contexte. Alors il existe* $u, v, x, y, z \in \Sigma^*$ *vérifiant*
- $uvxyz = w$*,*
- $|vxy| \leq p$*,*
- $vy \neq \varepsilon$*,*

- $uv^ixy^iz \in L_1$ *quel que soit* $i \in \mathbb{N} = \{0, 1, 2, \ldots\}$.

Puisque $|vxy| \leq p$, *le mot* vxy *ne peut avoir qu'une des formes suivantes:* a^k, a^sb^t, b^k, b^sc^t, c^k, *c'est-à-dire, au moins un des trois symboles* a, b, c *n'est pas dans* vxy. *Il est découle que aucun mot* uv^ixv^iz *n'est dans* L_1 *si* $i \neq 1$ *car le nombre d'au plus deux symboles peuvent changer. Ceci contredit le lemme de pompage et on conclut que* L_1 *n'est pas hors-contexte.*
Pour plus de précision, décortiquons les cas.

- $vxy = a^k$. *Alors* $u = a^\ell$, $v = a^n$, $x = a^m$, $y = a^r$ *et* $z = a^{p-\ell-n-m-r}b^pc^p$ *et pour* $i = 0$, $uv^ixv^iz = a^{\ell+m}a^{p-\ell-n-m-r}b^pc^p \notin L_2$ *(il y a* $n + r$ *a de moins que de* b, c*). Si* $uvx = b^k$ *ou* $uvx = c^k$, *l'argument est semblable.*

- $vxy = a^sb^t$. *Alors* v *et* y *contiennent chacun au plus une sorte de symbole (i.e.* $u = a^n$, $y = b^m$, uv^0xy^0z *contient moins de* a *ou de* b *(ou les deux) que de* c *et n'est pas dans* L_2. *Si un des deux contient deux sortes de symbole, le mot* uv^2xy^2z *ne sera pas de la bonne forme - si* $v = a^nb^m, n, m > 0$, *alors* $v^2 = a^nb^ma^nb^m$; *si* $y = a^nb^m$, *c'est analogue. L'argument est semblable pour* $vxy = b^sc^t$.

Exemple 4 *Avec le même* $\Sigma = \{a, b, c\}$, *soit* $L_2 = \{a^ib^jc^k : i < j < k\}$. *Si* L_2 *est hors-contexte, il existe une constante de pompage* p *et des mots* $u, v, x, y, z \in \Sigma^*$ *tel que* $a^pb^{p+1}c^{p+2} = uvxyz$, $|vxy| \leq p$, $|vy| > 0$ *et* $uv^nxy^nz \in L_2$ *pour* $n \in \mathbb{N}$. *On observe, comme avant, que dans ce cas* uvx *contient au plus deux symboles différents. Si* $uvx = a^k$, uv^2xy^2z *aura plus de* a *que de* b *(et c'est semblable si* $vxy = b^k$*). Si* $vxy = c^k$, v^0xy^0 *aura au plus autant de* c *que de* b. *Donc dans tous les cas "monosymboles" on sort du* L_2. *Si* $vxy = a^nb^m$, *soit* v *ou* y *contient deux sortes de symboles et alors* uv^2xy^2z *n'est pas de la bonne forme (des* b *suivis de* a*), soit* $v = a^n$, $y = b^m$ *et alors on a plusieurs possibilités. Si* $m > 0$, uv^2xy^2z *aura au moins autant de* b *que de* c. *Si* $m = 0$, *alors* uv^2xy^2z *aura au moins autant de* a *que de* b. *L'argument est semblable si* $vxy = b^nc^m$.