TD 9 : Hiérarchie de Chomsky

Notations

— Une grammaire *syntagmatique* (générale, de type 0, etc.) est un tuple $G = \langle V, \Sigma, P, S\rangle$ avec $V \cap \Sigma = \emptyset$, $P \subseteq (V \cup \Sigma)^* \times (V \cup \Sigma)^*$ fini et $S \in V$.

— Une telle grammaire est *monotone* si pour toute production $\alpha \to \beta$ de P, $|\alpha| \leq |\beta|$ (ou $\alpha = S$ et $\beta = \varepsilon$ et dans ce cas S n'apparaît dans la partie droite d'aucune règle).

— Elle est *contextuelle avec effacement* si toute règle de P est de la forme $\alpha_1 A \alpha_2 \to \alpha_1 \beta \alpha_2$ avec $\alpha_1, \alpha_2, \beta$ dans $(V \cup \Sigma)^*$ et A dans V.

— Elle est *contextuelle* (contextuelle sous forme normale, de type 1, etc.) si β est pris dans $(V \cup \Sigma)^+$ (ou $\alpha_1 = \alpha_2 = \beta = \varepsilon$ et $A = S$ et dans ce cas S n'apparaît dans la partie droite d'aucune règle).

— Elle est *algébrique* (non contextuelle, hors contexte, de type 2, etc.) si $P \subseteq V \times (V \cup \Sigma)^*$.

— Elle est *régulière* (de type 3, etc.) si elle est soit linéaire droite ($P \subseteq V \times \Sigma^* \cdot (V \cup \{\varepsilon\})$ soit linéaire gauche ($P \subseteq V \times (V \cup \{\varepsilon\}) \cdot \Sigma^*$).

Exercice 1 (Exemples de langages)**.**

1. Donner une grammaire monotone pour le langage $\{a^{n^2} \mid n \geq 0\}$.
2. Même question pour le langage $\{ww \mid w \in \{a, b\}^*\}$.

Exercice 2 (Équivalence entre grammaires)**.**

1. Montrer que pour toute grammaire syntagmatique (resp. monotone, contextuelle avec effacement, contextuelle) G, il existe une grammaire G' syntagmatique (resp. monotone, contextuelle avec effacement, contextuelle) équivalente telle que les règles de P' soient (de plus) de la forme

$$\alpha \to \beta \text{ avec } \alpha \in V^+ \text{ et } \beta \in V^*, \text{ ou de la forme}$$
$$A \to a \text{ avec } A \in V \text{ et } a \in \Sigma.$$

2. Montrer que pour toute grammaire monotone G, il existe une grammaire monotone G' équivalente telle que, si $\alpha \to \beta$ est une règle de P', alors $|\beta| \leq 2$ et $\alpha \in V^+$.
3. En déduire que pour toute grammaire monotone G, il existe une grammaire contextuelle G' équivalente.
4. Étendre la démonstration précédente pour montrer que pour toute grammaire syntagmatique G, il existe une grammaire contextuelle avec effacement G' équivalente.

Exercice 3 (Grammaires linéaires droites et gauches)**.** Montrer que pour toute grammaire linéaire droite, on peut construire une grammaire linéaire gauche équivalente. Faire de préférence une preuve directe, qui n'utilise pas la rationalité du langage.

Exercice 4 (Théorème de Chomsky et Schützenberger)**.** Le théorème dit qu'un langage L est algébrique si et seulement s'il existe un entier n, un langage rationnel R et un morphisme π tels que $L = \pi(D_n^* \cap R)$, où D_n^* dénote l'ensemble des mots bien parenthésés sur un alphabet à n paires de parenthèses.

L'intuition derrière ce théorème est que l'on peut séparer les aspects de structure (le parenthésage, c'est-à-dire la structure d'arbre) et de contrôle (le langage rationnel) d'un langage algébrique – ce dont on voit une autre interprétation avec les automates à pile.

1. Soit G une grammaire algébrique sur Σ. Proposer une grammaire algébrique G', qui explicite la structure des dérivations de G au moyen d'un alphabet Σ_n de n sortes de parenthèses, telle que
$$L(G') \subseteq D_n^* \text{ et } L(G) = \pi(L(G'))$$
avec π une projection de $\Sigma_n^* \to \Sigma^*$.
2. Il faut maintenant trouver un langage rationnel R tel que
$$L(G') = D_n^* \cap R\ .$$
Proposer un tel langage.
3. Montrer qu'il existe un morphisme μ de $\Sigma_n \to \Sigma_2$ tel que
$$D_n^* = \mu^{-1}(D_2^*)\ .$$
En déduire une autre formulation du théorème.