Université Paris Cité
Automates et Analyse Lexicale

UFR Informatique
Année 2023–2024

Handout 9

Le Lemme de l'Étoile

1 Le Lemme de l'Étoile

Aussi appelé *lemme d'itération*, en anglais *pumping lemma*.

1.1 L'énonce du lemme d'étoile

On dit qu'un langage $L \subseteq \Sigma^*$ a la propriété d'itération si :

il existe un entier $N \geq 1$
tel que pour tout mot $u \in L$ avec $|u| \geq N$
il existe des mots $x, y, z \in \Sigma^*$ avec $|xy| \leq N$, $|y| \geq 1$, $u = xyz$
tel que pour tout entier $k \geq 0$: $xy^kz \in L$.

Ce qu'on peut écrire dans le langage de la logique du premier ordre :

$$\exists N \geq 1 : \forall u \in L, |u| \geq N : \exists x, y, z \in \Sigma^*, |xy| \leq N, |y| \geq 1, u = xyz : \forall k \geq 0 : xy^kz \in L$$

Le lemme d'étoile dit : tout langage rationnel a la propriété d'itération.

Dans la preuve de ce lemme, la valeur N correspond au nombre d'états d'un automate qui reconnaît L.

Attention, ils peuvent exister des langages qui ont la propriété d'itération et qui ne sont pas rationnels.

1.2 Utiliser le lemme d'étoile pour montrer qu'un langage n'est *pas* rationnel

L'utilisation la plus importante du lemme d'étoile est pour montrer qu'un langage L n'est pas rationnel.

Soit L un langage. Si on arrive à montrer que L n'a pas la propriété d'itération alors on a une preuve que L n'est pas rationnel.

Pour montrer que L n'a *pas* la propriété d'itération il faut montrer la négation de la propriété d'itération, c'est à dire :

pour tout entier $N \geq 1$
il existe un mot $u \in L$ avec $|u| \geq N$
tel que pour tous mots $x, y, z \in \Sigma^*$ avec $|xy| \leq N$, $|y| \geq 1$, $u = xyz$
il existe un entier $k \geq 0$ tel que $xy^kz \notin L$.

ou encore en langage logique :

$$\forall N \geq 1 : \exists u \in L, |u| \geq N : \forall x, y, z \in \Sigma^*, |xy| \leq N, |y| \geq 1, u = xyz : \exists k \geq 0 : xy^kz \notin L$$

On peut voir une telle formule logique avec une alternance de quantificateurs $\forall$ et $\exists$ comme un jeu entre un joueur Existentiel qui peut choisir des valeurs des variables existentielles avec le but de montrer que ce qu'on à la fin de la formule est vrai, et le joueur Universel qui choisit les valeurs des variables universelles avec le but d'empêcher le jouer Existentiel. Donc, pour montrer qu'un langage L n'a pas la propriété d'itération :

1. Universel choisit un $N \geq 1$
2. Existentiel choisit un $u \in L$ avec $|u| \geq N$
3. Universel choisit un découpage $u = xyz$ avec $|xy| \leq N$, $|y| \geq 1$
4. Existentiel choisit $k \geq 0$

Si à la fin $xy^kz \notin L$ c'est Existentiel qui gagne la partie, sinon c'est Universel qui gagne. À chaque moment, les joueurs connaissent les choix précédents de leur opposant, et peuvent faire leur choix en conséquence.

Finalement, la formule est *vraie* si Existentiel a une *stratégie gagnante*, c'est à dire il a une stratégie de choisir les valeurs des variables existentiels qui le fait toujours gagner, peut importe les choix du joueur universel.

1.3 Utiliser les propriétés de clôture pour montrer qu'un langage n'est pas rationnel

Nous avons vu jusqu'à maintenant certaines propriétés de clôture de la classe des langages rationnels : sous union, intersection, complément, concaténation, étoile de Kleene. Il y a encore des autres que nous allons voir en cours ou en TD.

Si on connaît des langages qui ne sont pas rationnels on peut s'en servir pour montrer qu'un autre langage n'est pas rationnel.

Exemple : Si on a que $L_1 \cap L_2 = L_3$, et si on sait que L_2 est rationnel (par exemple quand il est donné par une expression rationnelle) et que L_3 n'est pas rationnel (par exemple parce qu'on a montré qu'il n'a pas la propriété d'itération) on peut conclure que L_1 n'est pas rationnel non plus.