Université Paris Cité
Automates et Analyse Lexicale

UFR Informatique
Année 2023–2024

# Handout 5

Automates avec $\epsilon$-transitions, Algorithme de Thompson

## 1 Automates avec $\epsilon$-transitions

Un *automate fini non déterministe avec $\epsilon$-transitions* (AFNDE) est un quintuplet $(\Sigma, Q, I, F, \delta)$ tel que

- $\Sigma$ est un alphabet fini ;
- $Q$ est un ensemble fini, appelé l'ensemble des *états* ;
- $I \subseteq Q$, appelé *ensemble des états initiaux* ;
- $F \subseteq Q$, appelé *ensemble des états acceptants* ;
- $\delta \colon Q \times \Sigma \cup \{\epsilon\} \mapsto \mathcal{P}(Q)$ est une fonction, appelée la *fonction de transition.*

La seule différence avec les AFND est qu'on a maintenant aussi des transitions qui ne consomment pas une lettre de $\Sigma$. Ces transitions permettent de se "téléporter" d'un état vers un autre.

Attention un AFNDE peut avoir des cycles de flèches qui sont toutes étiquetées par $\epsilon$, ce qui permet de tourner en rond sans d'avancer dans la lecture du mot.

### 1.1 $\epsilon$-clôture d'un ensemble d'états

L'$\epsilon$-clôture d'un ensemble d'états $P$ est l'ensemble de tous les états qu'on peut atteindre à partir d'un état de $P$ en un nombre quelconque d'$\epsilon$-transitions. On a en particulier pour tout $P \subseteq Q$, que $P$ est inclus dans son $\epsilon$-clôture.

Étant donné un AFNDE $(\Sigma, Q, I, F, \delta)$, et $P \subseteq Q$, l'ensemble $\epsilon$-*clôture*$(P)$ est

$$\{q \in Q \mid \text{existe } p_1, \ldots, p_n \in Q \text{ t.q. pour tout } 1 \leq i < n : p_{i+1} \in \delta(p_i, \epsilon), p_1 \in P, q = p_n\}$$

### 1.2 Le langage reconnu par un AFNDE

Étant donné un AFNDE $(\Sigma, Q, I, F, \delta)$, on définit une fonction $\delta^* \colon Q \times \Sigma^* \mapsto \mathcal{P}(Q)$ par récurrence sur le deuxième argument :

$$\begin{array}{rcl} \delta^*(q, \epsilon) & = & \{q\} \\ \delta^*(q, wa) & = & \displaystyle\bigcup_{p \in \delta^*(q,w)} \epsilon\text{-}\mathit{clôture}(\delta(p, a)) \end{array}$$

Le langage *reconnu* par l'AFNDE $A$ est

$$\mathcal{L}(A) = \{w \in \Sigma^* \mid \text{il existe } q_0 \in \epsilon\text{-}\mathit{clôture}(I) \text{ tel que } \delta^*(q_0, w) \cap F \neq \emptyset\}$$

L'AFNDE $A$ *accepte* un mot $w$ si $w \in \mathcal{L}(A)$.

## 2 Élimination des $\epsilon$-transitions

Pour tout AFNDE $A = (\Sigma, Q, I, F, \delta)$ il existe un AFND $A' = (\Sigma, Q', I', F', \delta')$ tel que $\mathcal{L}(A') = \mathcal{L}(A)$.

La construction de $A'$ est comme suit :

- $Q' = \{q \in Q \mid \text{existe } a \in \Sigma \text{ t.q. } \delta(q, a) \neq \emptyset\} \cup F$,
- $I' = \epsilon\text{-}\mathit{clôture}(I) \cap Q'$,
- $F' = F$,
- pour tout $q \in Q'$, $a \in \Sigma$ : $\delta'(q, a) = \epsilon\text{-}\mathit{clôture}(\delta(q, a)) \cap Q'$.

On a donc que pour tout langage $L$ les trois propriétés suivantes sont équivalentes :

1. $L$ est reconnu par un automate déterministe ;
2. $L$ est reconnu par un automate non déterministe ;
3. $L$ est reconnu par un automate non déterministe avec $\epsilon$-transitions.

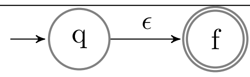
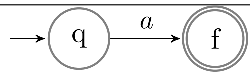
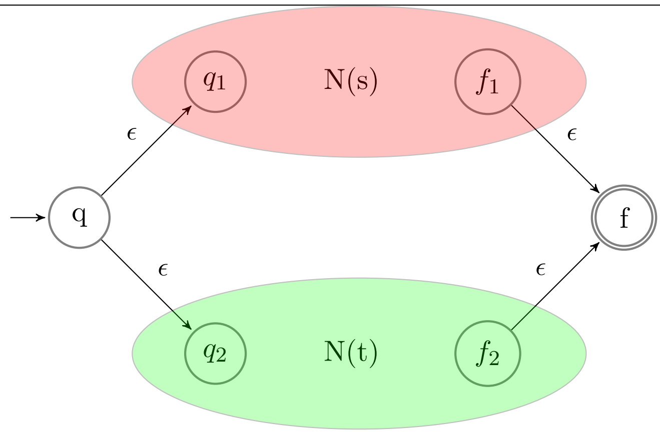
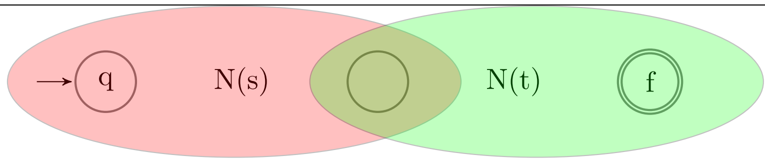
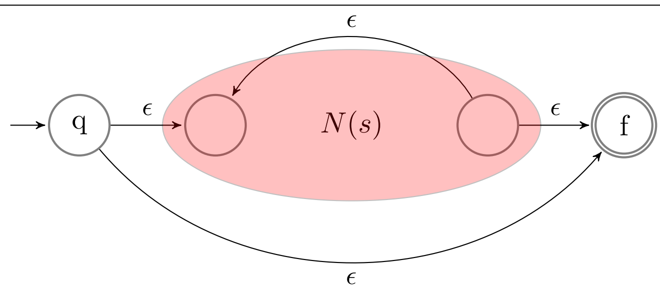
## 3 L'algorithme de Thompson

L'algorithme de Thompson transforme une expression rationnelle $r$ en un automate $A$ tel que $\mathcal{L}(r) = \mathcal{L}(A)$.

Dans un premier temps on construit un AFNDE par induction sur la structure de l'expression rationnelle. L'algorithme maintient un invariant, c-à-d tous les automates construits par cet algorithme satisfont ces propriétés :

- un seul état initial
- un seul état acceptant
- l'état initial n'est pas l'état acceptant
- aucune transition entre dans l'état initial
- aucune transition sort de l'état acceptant

Construction d'un AFNDE $N(s)$ pour une expression rationnelle $s$ :

| | |
|---|---|
| $N(\epsilon)$ |  |
| $N(a), a \in \Sigma$ |  |
| $N(\emptyset)$ |  |
| $N(s+t)$ |  |
| $N(st)$ |  |
| $N(s^*)$ |  |

Finalement on transforme cet AFNDE en un AFND (élimination des $\epsilon$-transitions), puis en un AFD (déterminisation).