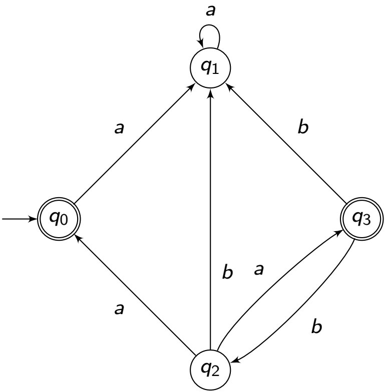
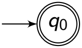
Langages et Automates
Résiduels et piles

Engel Lefaucheux

Prépas des INP

Minimiser un automate





Ces deux automates ont le même langage : $\{\varepsilon\}$

Quel algorithme pour minimiser un automate ?

Langage résiduel

Soit L un langage, $w \in \Sigma^*$.
Le résiduel de L par w est

$$w^{-1}L = \{v \in \Sigma^* \mid wv \in L\}$$

Exemple : $(aa)^{-1}((a+b)^2 b^* = b^*$

Exercice

Calculer

- $(aa)^{-}1(b^*ab^*ab^*)$
- $(\varepsilon)^{-}1(\varepsilon)$
- $(aa)^{-}1(b(a+b)^*)$

Calculer l'ensemble des résiduels possibles des langages

- a^*b^*
- $aa(ba)^*$

Automate des résiduels

Pour un langage régulier L sur Σ, on construit l'automate $\mathcal{A} = (Q, \Sigma, T, I, F)$ où

- $Q = \{q_{L'} \mid L' \text{ est un résiduel de } L\}$
- $I = \{q_L\}$
- $F = \{q_{L'} \mid \varepsilon \in L'\}$
- $(q_{L_1}, a, q_{L_2}) \in T$ ssi $L_2 = a^-1(L_1)$

Construire l'automate des résiduels de

- a^*b^*
- $aa(ba)^*$

Notez que cet automate est déterministe.

L'automate des résiduels est correct

Soit $w \in \Sigma^*$ et A l'automate des résiduels de L.

Montrons par récurrence sur $|w|$ que la lecture de w dans A termine dans l'état associé au résiduel $w^{-1}(L)$.

- si $|w| = 0$, alors $w = \varepsilon$, et le seul chemin étiquetté par ε est celui restant dans l'état initial: q_L et $L = \varepsilon^-1(L)$
- si $|w| = n > 0$, $w = ua$ où $u \in \Sigma^*$ et $a \in \Sigma$. Le chemin étiquetté par u finit dans l'état $q_{u^{-1}(L)}$ par hypothèse de récurrence. De plus, la seule transition étiquetté par a sortant de $q_{u^{-1}(L)}$ est $(q_{u^{-1}(L)}, a, q_{a^{-1}(u^{-1}(L))})$ et $a^{-1}(u^{-1=(L)}) = (ua)^{-1}(L) = w^{-1}(L)$ (voir DM)

Par ailleurs, un mot w est dans L ssi $\varepsilon \in w^{-1}(L)$.
Donc A accepte w ssi $w \in L$.

L'automate des résiduels est optimal

Soit L un langage, A l'automate des résiduels de L et A' un autre automate déterministe complet ayant moins d'états que A.

A' étant déterministe et complet, il existe une paire de mots w_1 et w_2 tels que $w_1^{-1}(L) \neq w_2^{-1}(L)$ et un état q de A' tel que les chemins étiquettés par w_1 et w_2 terminent en q.
$\rightarrow$ Il suffit de prendre un mot permettant d'accéder à chaque état de A, de voir dans quel état de A' ce mot termine, et d'appliquer le principe des tiroirs.

Soit $u \in w_1^{-1}(L) \setminus w_2^{-1}(L)$. On a donc que $w_1 u \in L$ et $w_2 u \notin L$. Hors, par déterminisme, A' accepte soit $w_1 u$ ET $w_2 u$ soit n'accepte ni $w_1 u$, ni $w_2 u$.

Donc A' ne reconnait pas L.

Alternative au lemme de l'étoile

Theorem (Théorème de MYHILL-NERODE)

Un langage est régulier si et seulement s'il possède un nombre fini de résiduels.

Un monde infini

- un système fini (les automates de ce cours) conserve une quantité d'information limitée
- l'environnement envoie une quantité d'information non bornée

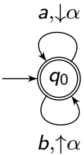
Comment retenir ces informations ?

Un monde infini

- un système fini (les automates de ce cours) conserve une quantité d'information limitée
- l'environnement envoie une quantité d'information non bornée

Comment retenir ces informations ?

Une possibilité : l'utilisation d'une pile



- $\downarrow\alpha$ ajoute α sur la pile
- $\uparrow\alpha$
  - ne peut se prendre que si α est au sommet de la pile
  - retire α du somme de la pile.

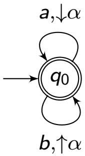
Quel langage pour cet automate ?

Un monde infini

- un système fini (les automates de ce cours) conserve une quantité d'information limitée
- l'environnement envoie une quantité d'information non bornée

Comment retenir ces informations ?

Une possibilité : l'utilisation d'une pile



- $\downarrow \alpha$ ajoute α sur la pile
- $\uparrow \alpha$
  - ne peut se prendre que si α est au sommet de la pile
  - retire α du somme de la pile.

Quel langage pour cet automate ?

$\{w \in \{a, b\}^* \mid \forall v, \text{prefix de } w, |v|_a \geq |v|_b\}$

Exercice

Construisez un automate à pile reconnaissant les langages suivants

- $\{a^n b^n \mid n \in \mathbb{N}\}$
- $\{a^n b^m c^m d^n \mid n, m \in \mathbb{N}\}$

Le langage suivant peut-il être reconnu par un automate à pile :
$\{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$

Langage hors contexte

Les langages reconnus par les automates à pile sont appelés langages hors-contexte.

Theorem (Théorème de l'étoile des automates à pile)

Soit L un langage hors contexte. Il existe $N \in \mathbb{N}$ tel que pour tout mot $w \in L$, si $|w| \geq N$, alors $w = uvwxy$ avec

- $|vwx| \leq N$
- $|vx| > 0$
- *pour tout $k \in \mathbb{N}$, $uv^nwx^ny \in L$*

Donc pour $\{a^nb^mc^nd^m \mid n, m \in \mathbb{N}\}$?

Langage hors contexte

Les langages reconnus par les automates à pile sont appelés langages hors-contexte.

Theorem (Théorème de l'étoile des automates à pile)

Soit L un langage hors contexte. Il existe $N \in \mathbb{N}$ tel que pour tout mot $w \in L$, si $|w| \geq N$, alors $w = uvwxy$ avec

- $|vwx| \leq N$
- $|vx| > 0$
- *pour tout* $k \in \mathbb{N}$, $uv^n wx^n y \in L$

Donc pour $\{a^n b^m c^n d^m \mid n, m \in \mathbb{N}\}$?

Les langages hors contexte correspondent aux langages générés par des grammaire de type 2.