LYCÉE LECONTE DE LISLE

Lemme d'Arden et applications

Vincent Picard

Lemme d'Arden

Soit A, B, L des langages vérifiant l'équation

$$L = AL \cup B$$

d'inconnue L alors

1. $L = A^*B$ est le plus petit langage solution de cette équation;
2. Si de plus $\varepsilon \notin A$, alors $L = A^*B$ est l'unique solution de cette équation.

- Ce lemme s'utilise pour résoudre des équations sur les langages.

Démonstration

- Soit A, B, L des langages vérifiant l'équation $L = AL \cup B$, montrons déjà que $L = A^*B$ est bien solution :

$$A(A^*B) \cup B = A\left(\bigcup_{n\in\mathbb{N}} A^n\right)B \cup B = \bigcup_{n\in\mathbb{N}^*} A^nB \cup B = \left(\bigcup_{n\in\mathbb{N}^*} A^n \cup \{\varepsilon\}\right)B = \bigcup_{n\in\mathbb{N}} A^nB = A^*B$$

- Soit A, B, L des langages vérifiant l'équation $L = AL \cup B$, montrons que $A^*B \subset L$: on montre par récurrence la propriété $P(k)$: "$A^kB \subset L$" ($k \in \mathbb{N}$).
  - **Initialisation** :

$$A^0B = \{\varepsilon\}B = B \subset AL \cup B = L$$

  - **Hérédité** : on suppose $P(k)$, on montre $P(k+1)$:

$$A^{k+1}B = AA^kB \subset AL \subset AL \cup B = L$$

  - **Conclusion** : Pour tout $k \in \mathbb{N}$, $P(k)$ est vrai et alors

$$A^*B = \left(\bigcup_{k\in\mathbb{N}} A^k \right) B = \bigcup_{k\in\mathbb{N}} \left(A^k B \right) \subset L$$

- Ajoutons l'hypothèse $\varepsilon \notin A$ et montrons qu'il n'y a pas d'autre solution que A^*B. Montrons par récurrence forte sur la longueur de u que pour tout mot $u \in L$, $u \in A^*B$.
  - **Initialisation** : $\varepsilon \in L = AL \cup B$, comme $\varepsilon \notin A$, cela signifie que $u \notin AL$, donc $u \in B$ donc $u \in A^*B$.
  - **Hérédité** : Soit $u \in L = AL \cup B$, on raisonne par disjonction de cas :
    - si $u \in B$, alors $u \in A^*B$,
    - si $u \in AL$ alors u s'écrit $u = vw$ avec $v \in A$ et $w \in L$. Comme $\varepsilon \notin A$, on a $|w| < |u|$ et par hypothèse de récurrence, $w \in A^*B$. On en déduit que $u = \underbrace{v}_{\in A} . \underbrace{w}_{\in A^*B} \in A^*B$.
  - **Conclusion :** On a montré par récurrence forte que $L \subset A^*B$ et ainsi $L = A^*B$.

Lien avec les mathématiques

- L'équation algébrique

$$x = ax + b$$

possède une unique solution lorsque $a \neq 1$:

$$x = \frac{1}{1-a} b$$

- L'hypothèse $\varepsilon \notin A$ est l'analogue de $a \neq 1$
- A^* est l'analogue de $\frac{1}{1-a}$, d'ailleurs remarquer la proximité de :
  - $A^* = \{\varepsilon\} \cup A \cup A^2 \cup A^3 \cup \ldots$
  - $1/(1-a) = 1 + a + a^2 + a^3 + \ldots$

Résolution d'équations sur les langages

- **Exceptionnellement, et uniquement dans ce cas, on notera + l'union de deux langages**

- Soit le système d'équations

$$\begin{cases} X = bX + abY & (E_1) \\ Y = aX + bY + \varepsilon & (E_2) \end{cases}$$

où X et Y sont les langages inconnus.

- **Résolution**

  - On applique le lemme d'Arden dans (E_2) avec l'inconnue Y. On trouve :

$$Y = b^*(aX + \varepsilon) = b^*aX + b^*$$

  - On substitue Y par son expression en fonction de X dans (E_1) :

$$X = bX + abb^*aX + abb^* \quad (E_1)$$

- On applique une nouvelle fois le lemme d'Arden dans (E_1) avec l'inconnue X :

$$X = (b + abb^*a)^* abb^*$$

- On retrouve la valeur de Y à partir de celle de X :

$$Y = b^*aX + b^* = b^*a(b + abb^*a)^* abb^* + b^*$$

- **Conclusion :** l'unique couple (X, Y) solution est :

$$\begin{cases} X = (b + abb^*a)^* abb^* \\ Y = b^*a(b + abb^*a)^* abb^* + b^* \end{cases}$$

- La méthode se généralise avec plus d'inconnues et d'équations : on utilise la dernière équation pour obtenir une expression de la dernière inconnue en fonction des autres, on substitue dans les autres équations et on obtient un système avec une équation et une inconnue de moins, et ainsi de suite...

Exercices

- Résoudre les systèmes d'équations suivants :

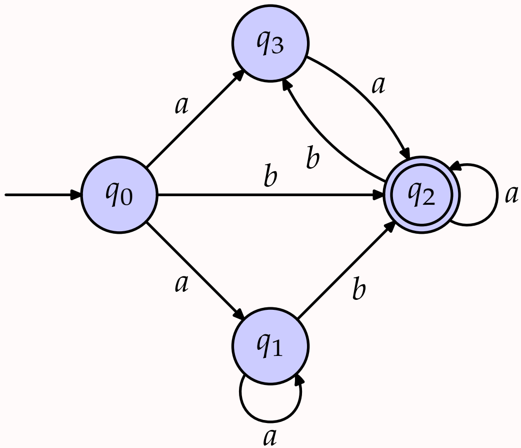
$$\begin{cases} X = bY + aZ \\ Y = aX + bZ + \varepsilon \\ Z = aZ \end{cases}$$

$$\begin{cases} X_1 = bX_2 + aX_4 \\ X_2 = aX_3 + bX_4 \\ X_3 = aX_4 + bX_5 + \varepsilon \\ X_4 = aX_4 \\ X_5 = aX_3 + bX_4 \end{cases}$$

- Remarquer qu'en pratique, on applique le lemme d'Arden le plus tard possible.

Des automates vers les expressions régulières : méthode par équations

- On note L_i le langage des mots reconnus à partir de l'état q_i
- **But :** déterminer L_0



$$\begin{cases} L_0 = aL_1 + bL_2 + aL_3 & (E_1) \\ L_1 = aL_1 + bL_2 & (E_2) \\ L_2 = aL_2 + bL_3 + \varepsilon & (E_3) \\ L_3 = aL_2 & (E_4) \end{cases}$$

Résolution

- L_3 est déjà déterminée par (E_4), on substitue dans les autres équations :

$$\begin{cases} L_0 = aL_1 + (b + a^2)L_2 & (E_1) \\ L_1 = aL_1 + bL_2 & (E_2) \\ L_2 = (a + ba)L_2 + \varepsilon & (E_3) \end{cases}$$

- On applique le lemme d'Arden dans (E_3), pour déterminer L_2 :

$$L_2 = (a + ba)^*$$

- On substitue dans (E_1) et (E_2) :

$$\begin{cases} L_0 = aL_1 + (b + a^2)(a + ba)^* & (E_1) \\ L_1 = aL_1 + b(a + ba)^* & (E_2) \end{cases}$$

- On applique le lemme d'Arden dans (E_2) pour déterminer L_1 :

$$L_1 = a^*b(a + ba)^*$$

- On substitue dans (E_1) :

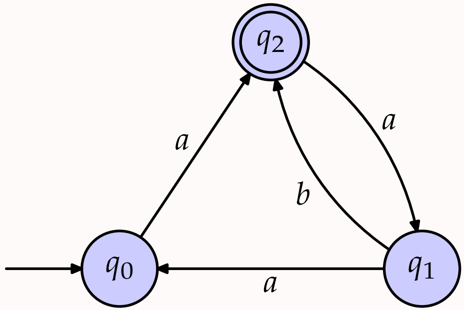
$$L_0 = aa^*b(a + ba)^* + (b + a^2)(a + ba)^* = (aa^*b + b + a^2)(a + ba)^*$$

- **Conclusion :** une expression régulière dénotant le langage reconnu par l'automate est :

$$\boxed{\left(aa^*b \mid b \mid a^2\right)(a \mid ba)^*}$$

Exercices

En utilisant le lemme d'Arden, déterminer une expression régulière dénotant le langage reconnu par les automates suivants :

1. 

2.