Les langages reconnus par les automates d'états finis sont facilement décrits par des expressions régulières. Elles ont été proposées par Kleene en 1956.

1. Langage rationnel et expression régulière

1.1 Définition: Soit X un alphabet, les expressions régulières sur X et les ensembles qu'elles dénotent sont définis récursivement comme suit :

1. $\varnothing$ est une expression régulière, $L = \varnothing$ représente l'ensemble vide
2. ε est une expression régulière, $L = \{\varepsilon\}$
3. $\forall w_i \in X$, w_i est une expression régulière, $L = \{w_i\}$
4. Si E_1 et E_2 sont deux expressions régulières alors :
   $E_1 \cup E_2$ est une expression régulière, $L(E_1 \cup E_2)= L(E_1) \cup L(E_2)$
   $E_1.E_2$ est une expression régulière, $L(E_1.E_2)= L(E_1).L(E_2)$
   E_1^* est une expression régulière.

1.2 Définition: Soit X un alphabet, L est un langage rationnel sur X si et seulement si, il peut-être exprimé en fonction des opérations : union, itération et concaténation.

1.3 Définition : Deux expressions régulières sont équivalentes si et seulement si, elles décrivent le même langage.

1.3 Priorité des opérations : L'opération la plus prioritaire est l'itération (*), suivie de la concaténation (**.**) et enfin de l'union ($\cup$).

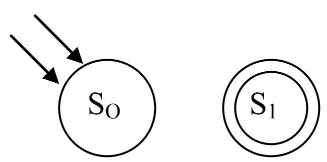
2. Propriétés

2.1 Théorème de Kleene : La classe des langages rationnels est égale exactement à la classe des langages réguliers.

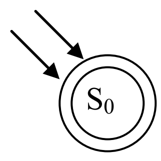
2.2 Proposition 1: Soit E une expression régulière, il existe un automate d'états finis généralisé qui accepte L(E).

Démonstration par récurrence sur n, le nombre d'opérations dans l'expression régulière :
a. Cas particuliers : n=0

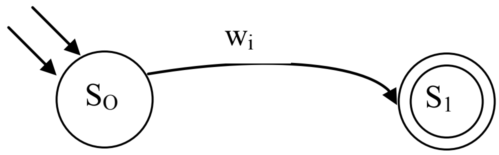
$\varnothing$ est une expression régulière,



ε est une expression régulière, $L = \{\varepsilon\}$



$\forall w_i \in X$, w_i est une expression régulière, $L = \{w_i\}$



b. Hypothèse de récurrence

Soient E_1 et E_2, deux expressions régulières, E_1 contenant n_1 opérations et E_2 contenant n_2. Soient A_1<X_1, S_1, S_{01}, F_1, II_1> et A_2<X_2, S_2, S_{02}, F_2, II_2> deux automates d'états finis tels que $L(E_1) = L(A_1)$ et $L(E_2) = L(A_2)$, $S_1 \cap S_2 = \varnothing$.

c. Montrons que :

- $E_1 \cup E_2$ est une expression régulière pour laquelle il existe un automate d'état finis A<X^*, S, S_0, F, II> tel que $L(A) = L(E_1) \cup L(E_2)$
  $X = X_1 \cup X_2$, $S = S_1 \cup S_2 \cup \{S_0\}$, $F = F_1 \cup F_2$, $II' = II \cup \{(S_0, \varepsilon, S_{01}), (S_0, \varepsilon, S_{01})\}$,
  Montrons que $L(A) = L(E_1) \cup L(E_2)$

- $E_1.E_2$ est une expression régulière pour laquelle il existe un automate d'état finis A<X^*, S, S_0, F, II> tel que $L(A) = L(E_1).L(E_2)$
  $X = X_1 \cup X_2$, $S = S_1 \cup S_2$, $S_0 = S_{01}$, $F = F_2$, $II' = II \cup \{(S_k, \varepsilon, S_{02}), \forall S_k \in F_1$.
  Montrons que $L(A) = L(E_1).L(E_2)$ (La démonstration a été faite en cours)

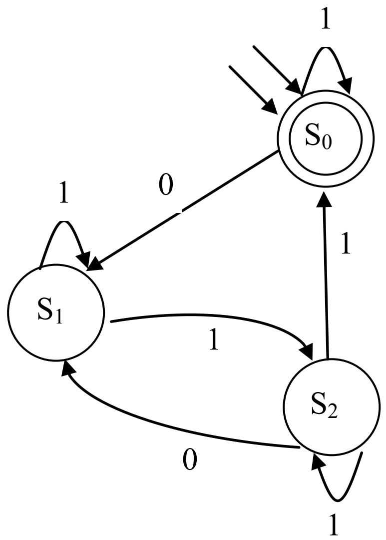
- E_1^* est une expression régulière pour laquelle il existe un automate d'état finis A<X^*, S, S_0, F, II> tel que $L(A) = L(E_1^*)$
  $X=X_1$, $S=S_1 \cup \{S_0, S_f\}$, $F=\{S_f\}$, $II' = II \cup \{(S_0, \varepsilon, S_{01}), (S_0, \varepsilon, S_f)\} \cup \{(S_k, \varepsilon, S_{01}), \forall S_k \in F_1\}$.
  Montrons que $L(A) = L(E_1^*)$.
  Démonstration laissée en exercice.

2.3 Proposition 2: A tout automate d'état finis A<X, S, S_0, F, II>, il existe une expression régulière E tel que $L(A) = L(E)$.

Démonstration :

Nous allons présenter les étapes de construction d'une expression régulière à partir d'un automate simple à travers un exemple, puis nous généraliserons les étapes.

Exemple :



1. Nous allons associer à chaque état de l'automate, une expression régulière qui permet de savoir comment les mots, qui font passer l'automate de cet état aux états finaux, s'écrivent. Nous nous intéresserons aux mots qui font passer l'automate de l'état initial à tous les états finaux.

Par exemple, à l'état S_1, nous allons associer l'expression régulière E_1. A partir de S_1, et à la lecture d'un 1, l'automate peut passer à l'état S_2 ou revenir à l'état S_1. Les mots que l'automate lit à partir de S_1 commencent forcement par un 1. L'expression régulière associé à S_1 est la suivante $E_1 = 1\ E_1 \cup 1\ E_2$.

Les expressions régulières E_2 et E_0, associées respectivement à S_2 et S_0 sont définies comme suit :

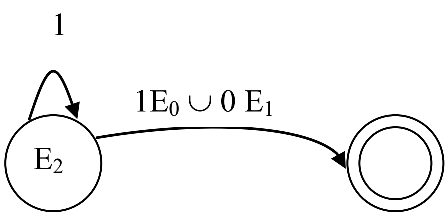
- $E_2 = 1\ E_0 \cup 0E_1 \cup 1E_2$
- $E_0 = 1\ E_0 \cup 0E_1 \cup \varepsilon$, le mot vide est rajouté car S_0 est un état final.

Nous avons établi trois équations à trois inconnus. Maintenant, nous allons résoudre ces équations en éliminant étape par étape les inconnus par une série de remplacement. L'expression régulière qui nous intéresse est celle de E_0. Dans cet exemple, l'état initial est aussi l'état final.

$$\begin{cases} E_0 = 1\ E_0 \cup 0E_1 \cup \varepsilon \\ E_1 = 1\ E_1 \cup 1\ E_2. \\ E_2 = 1\ E_0 \cup 0E_1 \cup 1E_2 \end{cases}$$

2. Elimination de l'inconnue E_2 :

$E_2 = 1E_0 \cup 0E \cup 1E_2$, on peut associer à cette expression régulière, l'automate suivant :



$E_2 = 1^* (1E_0 \cup 0E_1)$, (Application de la règle d'Arden)

Les mots à partir de S_2 s'écrivent comme une série de 1 (0, 1 ou plusieurs), suivi d'un 1 et le passage vers l'état S_0, ou suivi d'un 0 et le passage vers l'état S_1.

3. Remplacement de E_2 dans E_1 et élimination de E_1:

$E_1 = 1E_1 \cup 1[1^* (1E_0 \cup 0E_1)] = 11^*1E_0 \cup (1 \cup 11^*0)\ E_1$
$E_1 = (1 \cup 11^*0)^*\ 11^*1E_0$

4. Remplacement de E_1 dans E_0 et élimination de l'inconnue E_0 :

$E_0 = 1\ E_0 \cup 0E_1 \cup \varepsilon = 1E_0 \cup 0(1 \cup 11^*0)^*\ 11^*1E_0 \cup \varepsilon$
$E_0 = (1 \cup 0(1 \cup 11^*0)^*\ 11^*1)E_0 \cup \varepsilon$
$E_0 = (1 \cup 0(1 \cup 11^*0)^*\ 11^*1)^*.\ \varepsilon = (1 \cup 0(1 \cup 11^*0)^*\ 11^*1)^*$

Généralisation :

- On associe à chaque état, une expression régulière. Il y a autant d'expressions régulières qu'il y a d'états, soit n ce nombre.

$$E_k = \bigcup_{j=1}^{n} V_{kj}E_J \cup w_k \text{ avec } w_k = \begin{cases} \varepsilon \text{ si } S_k \in F \\ \varnothing \text{ Sinon} \end{cases} \qquad 1 \leq k \leq n$$

Avec $V_{kj} = \{x_i \in X \text{ tq } (S_k, x_i, S_J) \in II\}$

- Elimination des inconnues, étape par étape (application de la règle d'Arden):

$$E_1 = \bigcup_{j=1}^{n} V_{1J}E_J \cup w_1 = V_{11}\, E_1 \bigcup_{j=2}^{n} V_{1J}E_J \cup w_1$$

$$E_1 = V_{11}^{*} \left(\bigcup_{j=2}^{n} V_{1J}E_J \cup w_1\right)$$

3. Construction d'un automate à partir d'une expression régulière :

3.1 Définition : On définit sur les expressions régulières une opération appelée dérivée comme suit : Soit E une expression régulière on a $E//u = \{w \in X^* \,/\, uw \in\}$, u est un mot de X^*.

3.2 Exemple : $E = ab^*(a \cup b)^*$, $E//a = b^*(a \cup b)^*$,
$E//b = \varnothing$, aucun mot de L(E) ne commence par un b.

3.3 Proposition : Soient E_1 et E_2, deux expressions régulières on a :

1. $(E_1 \cup E_2)//u = (E_1//u) \cup (E_2//u)$, $u \in X^*$
2. $(E_1 \,.\, E_2)//u_i = (E_1//u_i) \,.\, E_2$, $u_i \in X$ si $\varepsilon \notin L(E_1)$
   $= (E_1//u_i) \,.\, E_2 \cup (E_2//u_i)$, $u_i \in X$ si $\varepsilon \in L(E_1)$
3. $E_1^*// u_i = (E_1// u_i) \,.\, E_1^*$
4. $E_1//(u.v)= (E_1//u)//v$.

3.4 Propriétés : Soient E_1, E_2 et E_3, trois expressions régulières on a :

1. $E_1 \cup E_2 = E_2 \cup E_1$
2. $E_1 \cup (E_2 \cup E_3) = (E_1 \cup E_2) \cup E_3$
3. $E_1 \,.\, (E_2 \,.\, E_3) = (E_1 \,.\, E_2) \,.\, E_3$
4. $E_1 \,.\, (E_2 \cup E_3) = (E_1 \,.\, E_2) \cup (E_1 \,.\, E_3)$
5. $(E_1 \cup E_2) \,.\, E_3 = (E_1 \,.\, E_3) \cup (E_2 \,.\, E_3)$
6. $E_1 \cup \varnothing = \varnothing \cup E_1$
7. $E_1 \,.\, \varepsilon = \varepsilon \,.\, E_1 = E_1$
8. $E_1 \,.\, \varnothing = \varnothing \,.\, E_1 = \varnothing$
9. $E_1 \cup E_1 = E_1$
10. $E_1^* \,.\, E_1^* = E_1^*$
11. $E_1 \,.\, E_1^* = E_1^* \,.\, E_1 = E_1^+$
12. $(E_1^* \cup E_2^*)^* = (E_1 \cup E_2)^* = (E_1^* \,.\, E_2^*)^*$
13. $(E_1 \cup E_2)^* = E_1^* \,.\, (E_2 \,.\, E_1^*)^*$

3.5 Proposition : A toute expression régulière E, il existe un automate d'états finis A<X,S,S_0,F,II> simple déterministe complet tq L(E)= L(A).

Nous allons montrer les étapes de construction de l'automate d'états finis avec la méthode des dérivées à travers un exemple, puis nous démontrerons la proposition:

Soit $E_0 = a^* b^*$

1. On associe à l'expression régulière initiale E_0, l'état initial S_0
2. On dérive E_0 par rapport à toutes les lettres de X, des états sont associés aux expressions régulières obtenues :

$E_0//a = (a^* b^*) //a = (a^* // a).b^* \cup b^*//a$ (proposition 1)
$= (a//a).a^* b^* \cup \emptyset = a^* b^* = E_0$ (proposition 3)

$E_0//a = (a^* b^*) //b = (a^* // b).b^* \cup b^*//b$ (proposition 1)
$= \emptyset \cup (b//b). b^* = b^* = E_1$ (proposition 3)

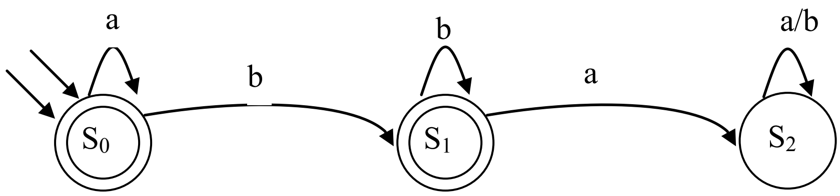
$E_1//a = E_0//aa = b^*//a = \emptyset = E_2$

$E_1//b = E_0//ab = b^*//b = b^* = E_1$

$E_2//a = E_0//aaa = \emptyset// a = \emptyset = E_2$ (état puit)

$E_2//b = E_0//aab = \emptyset//b = \emptyset = E_2$ (état puit)

Il n'y a plus de nouvelles expressions régulières générées. A chaque expression régulière est associée un état de l'automate. Une dérivée par rapport à une lettre de X représente l'étiquette d'une transition. L'automate qui reconnaît les mots dénotés par E_0 est le suivant :



ε appartient à $L(E_0)$ à $L(E_1)$, les états S_0 et S_1 deviennent donc des états finaux.

Demonstration de la proposition:

1. L'automate A<X, S, S_0, F, II> obtenu par la méthode des dérivées est un automate simple, déterministe et complet. Nous allons maintenant déterminer les paramètres de A :

X est l'alphabet utilisé par E,
$S = \{E_0//w, w \in X^*\}= \{E_j//w_i, w_i \in X\}$
$S_0 = E_0//\varepsilon = E_0$
$F = \{E_j//w_i, w_i \in X \text{ tq } \varepsilon \in L(E_j//w_i)\}$
$II : S \times X \rightarrow S$
$E_i,w_i \rightarrow E_j = E_i//w_i$

2. montrons que chaque dérivée correspond à une lecture dans l'automate. Cette démonstration se fait par récurrence sur le nombre de dérivation. Soit n ce nombre :

a. Cas particulier : n = 1

$E_i \underset{A}{\overset{w_i}{\vdash\!\!-}} E_j$ avec $w_i \in X$, par définition $E_j = E_i // w_i \in II$

b. Hypothèse de récurrence

$E_i \underset{A}{\overset{w}{\vdash\!\!-}} E_j$ avec $w \in X^*$, on a $E_j = E_i // w$

c. Montrons que la relation est toujours vraie à l'ordre supérieur

$E_i \underset{A}{\overset{w}{\vdash\!\!-}} E_j \underset{A}{\overset{x_i}{\vdash\!\!-}} E_k$ avec $w \in X^*$ et $x_i \in X$

$E_k = (E_J // x_i)$ cas particulier

$E_J = E_i // w$ Hypothèse de récurrence

$E_k = (E_i // w) // x_i$

$E_J = E_i // (wx_i)$

3. Montrons que $L(A) = L(E)$

$\forall w \in L(A) \Leftrightarrow S_0 \underset{A}{\overset{w}{\vdash\!\!-}} S_f \quad S_f \in F \Leftrightarrow \varepsilon \in L//w \Leftrightarrow w \in L$