Teoría de Autómatas y Lenguajes Formales

Boletín de Autoevaluación 4: **¿Cómo se calcula la Expresión Regular asociada a un AFD?.**

1. Objetivos.

El objetivo de este boletín es ilustrar uno de los métodos que permiten calcular la expresión regular que denota el mismo lenguaje que un Autómata Finito reconoce mediante ejemplos y, además, proporcionar la solución a alguno de los problemas propuestos en el boletín para que podáis comprobar si habéis aplicado bien este método.

2. Idea Principal.

Los teoremas de Kleene, el de Análisis y el de Síntesis, establecen la equivalencia entre los AFs y las expresiones regulares. En concreto, el de Análisis permite afirmar que dado un AFD, existe una expresión regular que lo denota. Su demostración ya establece un método para calcular dicha expresión regular, pero en la práctica no es el método que se suele utilizar para realizar dicho cálculo.

En su lugar, es más aconsejable el uso de la **Regla de Arden**. Para aplicarla, hay que establecer un sistema de ecuaciones lineales en expresiones regulares y resolverlo. Cada ecuación de un sistema de ecuaciones lineales en expresiones regulares tiene la siguiente forma general

$$X = rX + s$$

en la que r y s son expresiones regulares sobre un alfabeto Σ. Dependiendo de que $\lambda \in r$ o $\lambda \notin r$ se tienen dos soluciones para dicha ecuación:

- $\lambda \notin \mathbf{r}$ y, entonces, $X = r^*s$. Para comprobarlo, basta ver que

$$X = rX + s = r(r^*s) + s = (rr^* + \lambda)s = r^*s = X.$$

- $\lambda \in \mathbf{r}$ y, entonces, hay infinitas soluciones $X = r^*(s + t)$, donde t es una expresión regular cualquiera sobre Σ (normalmente, este caso no nos afectará). Este resultado puede comprobarse de la siguiente forma,

$$X = rX + s = r(r^*(s + t)) + s = (rr^* + \lambda)s + rr^*t = (r^+ + \lambda)s + r^+t.$$

Como $r^* = r^+$ si $\lambda \in r$, entonces lo anterior queda,

$$(r^+ + \lambda)s + r^+t = (r^* + \lambda)s + r^*t = r^*(s + t) = X.$$

3. Método para obtener la Expresión regular que denota a un AF dado.

El método se basa en el establecimiento de un sistema de ecuaciones lineales en expresiones regulares a partir del AF. Sea A,

$$A = \langle \Sigma, Q, f, q_0, F \rangle$$

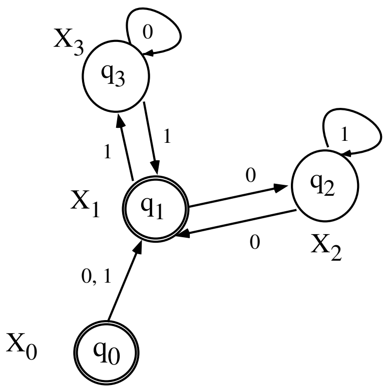
un AFD o un AFN; a A se le puede asociar un sistema de ecuaciones lineales en expresiones regulares de la siguiente forma:

1. Se asocia una variable a cada estado: $\forall q_i \in Q$, se le asocia X_i.
2. Las ecuaciones se construyen en función de las transiciones: si $q_j \in f(q_i, a)$, entonces en la ecuación de la variable X_i aparece el término aX_j en su parte derecha: $X_i = ... + aX_j + ...$.
3. Además, se asocia el término λ a los estados finales: si $q_i \in F$, entonces en la ecuación de la variable X_i aparece λ en su parte derecha: $X_i = ... + \lambda + ...$.

El lenguaje reconocido por el AF es la *expresión regular de la variable asociada a su estado inicial*.

3.1. Ejemplo 1.

Hay que calcular la expresión regular que denota el lenguaje reconocido por el autómata que se muestra en la siguiente figura:



Sobre la figura se ha asociado ya a cada estado su correspondiente variable. De acuerdo al segundo y tercer paso del método, el sistema de ecuaciones lineales que hay que resolver es el siguiente:

$$
\begin{aligned}
&X_0 = 0X_1 + 1X_1 + \lambda = (0+1)X_1 + \lambda \\
&X_1 = 0X_2 + 1X_3 + \lambda \\
&X_2 = 1X_2 + 0X_1 \\
&X_3 = 0X_3 + 1X_1
\end{aligned}
$$

Recordemos que la solución de la ecuación general, $X = rX + s$ es $X = r^*s$. En la última ecuación, si identificamos términos con r y s se tiene:

$$X_3 = \underbrace{0}_{r} X_3 + \underbrace{1X_1}_{s}.$$

Por lo tanto,

$$X_3 = 0^*1X_1.$$

De forma similar, se puede operar con X_2,

$$X_2 = \underbrace{1}_{r} X_2 + \underbrace{0X_1}_{s}, \Rightarrow X_2 = 1^*0X_1.$$

Se tienen X_2 y X_3 en función de X_1. Al substituir en la ecuación de X_1 se obtiene

$$X_1 = 0X_2 + 1X_3 + \lambda = 01^*0X_1 + 10^*1X_1 + \lambda = (01^*0 + 10^*1)X_1 + \lambda$$

Agrupando términos se llega a

$$X_1 = \underbrace{(01^*0 + 10^*1)}_{r} X_1 + \underbrace{\lambda}_{s}, \Rightarrow X_1 = (01^*0 + 10^*1)^*\lambda = (01^*0 + 10^*1)^*.$$

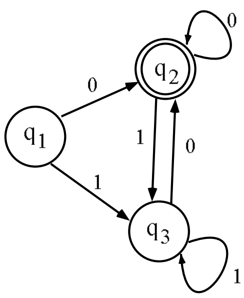
y, al substituir en la ecuación de X_0 se obtiene finalmente,

$$X_0 = (0+1)X_1 + \lambda = (0+1)(01^*0 + 10^*1)^* + \lambda,$$

$$\Rightarrow L(AFD) = (0+1)(01^*0 + 10^*1)^* + \lambda.$$

3.2. Ejemplo 2.

El siguiente ejemplo es el mismo que tenéis en los apuntes de teoría ilustrando el Teorema de Análisis (y así podréis comparar ambos métodos). El autómata es el siguiente:



Se asocia X_1 al estado q_1, X_2 al estado q_2 y X_3 al estado q_3. El sistema de ecuaciones lineales resultante es:

$$\begin{aligned} X_1 &= 0X_2 + 1X_3 \\ X_2 &= 0X_2 + 1X_3 + \lambda \\ X_3 &= 1X_3 + 0X_2 \end{aligned}$$

Podemos comenzar con la ecuación de X_3:

$$X_3 = \underbrace{1}_{r} X_3 + \underbrace{0X_2}_{s}, \Rightarrow X_3 = 1^*0X_2.$$

Se substituye la expresión obtenida en la ecuación de X_2,

$$X_2 = 0X_2 + 1X_3 + \lambda = 0X_2 + 11^*0X_2 + \lambda = \underbrace{0 + 11^*0}_{r} X_2 + \underbrace{\lambda}_{s}, \Rightarrow X_2 = (0 + 11^*0)^*.$$

Ahora se substituyen X_2 y X_3 en la ecuación de X_1:

$$X_1 = 0X_2 + 1X_3 = 0(0 + 11^*0)^* + 11^*0X_2 = 0(0 + 11^*0)^* + 11^*0(0 + 11^*0)^*$$

y, si se agrupan términos, se llega a

$$X_1 = (0 + 11^*0)(0 + 11^*0)^* = ((\lambda + 11^*)0)((\lambda + 11^*)0)^* = 1^*0(1^*0)^* = (1^*0)^*1^*0 = (1 + 0)^*0.$$

Es decir,

$$L(AFD) = (1 + 0)^*0.$$

Nota: Las expresiones regulares tienen un montón de propiedades y, a veces, nos asusta el "mogollón". En el desarrollo anterior hemos usado tres de las que más salen; recordadlas, que son muy útiles:

$$(1)\ (\lambda + \alpha\alpha^*) = \alpha^*,\ (2)\ \alpha\alpha^* = \alpha^*\alpha\ \ y\ \ (3)\ (\alpha^*\beta)^*\alpha^* = (\alpha + \beta)^*.$$

Otra buena política para evitar "enmogollanarse" es simplificar desde el principio todo lo que podamos. Del sistema original,

$$\begin{aligned} X_1 &= 0X_2 + 1X_3 \\ X_2 &= 0X_2 + 1X_3 + \lambda \\ X_3 &= 1X_3 + 0X_2 \end{aligned}$$

se sigue que $X_2 = X_1 + \lambda$. Si se usa esa relación, el resultado es bastante más simple:

$$X_3 = 1^*0X_2, X_2 = X_1 + \lambda \Rightarrow X_3 = 1^*0(X_1 + \lambda).$$

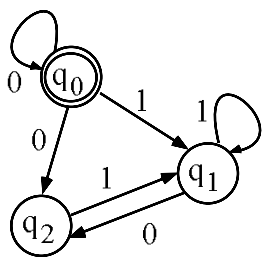
Al substituir en X_1:

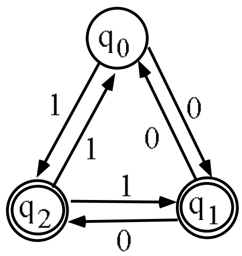
$$X_1 = 0X_2 + 1X_3 = 0(X_1 + \lambda) + 11^*0(X_1 + \lambda) = (0 + 11^*0)X_1 + (0 + 11^*0) = (0 + 11^*0)^*(0 + 11^*0)$$

se llega con menos pasos al mismo resultado (bueno, falta volver a aplicar las tres propiedades anteriores para obtener la expresión simplificada, pero... aplicadlas vosotros y vais entrenándoos ;-).

3.3. Autoevaluación.

1. Construir expresiones regulares que denoten los mismos lenguajes que los reconocidos por los siguientes autómatas finitos:





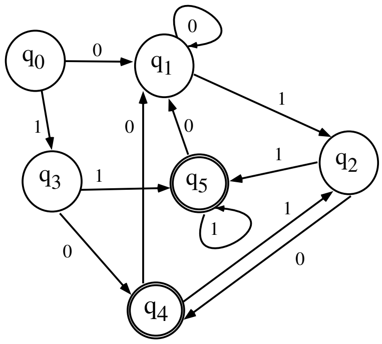
Solución: El primer autómata queda denotado por la expresión regular 0^*. La verdad es que se ve a "ojo"... en este problema lo complicado es obtenerla formalmente. Como pistas, las siguientes:

- la expresión regular $\emptyset$ concatenada con cualquier otra expresión regular es $\emptyset$ ($r\emptyset = \emptyset r = \emptyset,\ \forall\ expr.\ reg.\ r$: si no concateno nada con r ¿qué obtengo? ¡nada!).
- Por definición $\emptyset^* = \{\lambda, \emptyset\} = \lambda$.

Para acabar de liar el problema, tenéis el segundo autómata:

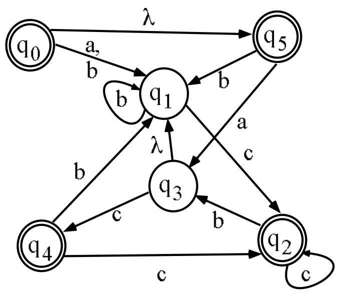
$$L(AFD) = ((0(10)^*(0+11) + 1(01)^*(1+00))^*(0(10)^*(1+\lambda) + 1(01)^*(0+\lambda))).$$

2. Calcular la expresión regular asociada al AFD mínimo que reconoce el mismo lenguaje que el representado en la figura (obtener antes el AFD mínimo).



Solución: AFD mínimo: $q_0 \equiv q_1$ y $q_2 \equiv q_3$. Expr. Reg.: $L(AFD) = (0+1)^*1(0+1)$

3. En el Boletín de Autoevaluación 1, se propuso la transformación del siguiente AFλ a AFD:



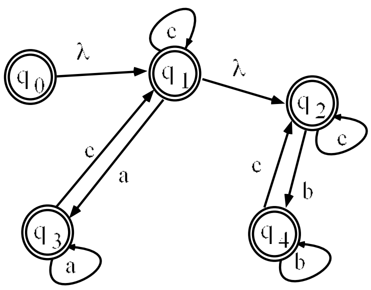
AFD equivalente:

	f	a	b	c
F	$\{q_0\}$	$\{q_1, q_3\}$	$\{q_1\}$	$\emptyset$
	$\{q_1, q_3\}$	$\emptyset$	$\{q_1\}$	$\{q_2, q_4\}$
	$\{q_1\}$	$\emptyset$	$\{q_1\}$	$\{q_2\}$
F	$\{q_2, q_4\}$	$\emptyset$	$\{q_1, q_3\}$	$\{q_2\}$
F	$\{q_2\}$	$\emptyset$	$\{q_1, q_3\}$	$\{q_2\}$

Completar el problema dando la expresión regular del AFD mínimo equivalente.

Solución: $L(AFD) = (a+b)(b+c)^*c + \lambda$.

4. Obtener el AFD mínimo que reconozca el lenguaje reconocido por el siguiente AFλ,



Calcular la expresión regular asociada al lenguaje reconocido por el autómata anterior.

Solución: Pistas: el AFD mínimo tiene 3 estados y los 3 son finales.

$L(AFD) = (a^*c)^*(a^* + b(b + c)^*)$.