Aplicación del Análisis de Kleene

Teoría de Lenguajes

2022

1 Introducción

Dado un **autómata finito determinista** M se quiere hallar una **expresión regular** r tal que $L(M) = L(r)$. En el curso vemos tres métodos que nos permiten eso:

1. Cálculo de R_{ij}^k
2. Análisis de Kleene
3. Cálculo de las clases de equivalencia de R_M

En este documento veremos cómo hallar una expresión regular r aplicando el **Análisis de Kleene**.

1.1 Análisis de Kleene

Este método intenta formar un sistema de ecuaciones donde se tiene una variable X_i por cada estado del autómata M. Es decir, que si el autómata cuenta con cuatro estados, el sistema de ecuaciones tendrá a X_0, a X_1, a X_2 y a X_3.

Para formular estos sistemas de ecuaciones el método intenta responder a la pregunta:

¿Cuáles son las tiras que, al ser reconocidas desde q_i, terminan en *algún* estado final?

Por esa razón, este método se fijará en las transiciones **salientes** y la expresión regular r buscada será la hallada para X_0[1].

Por más detalle en el enunciado de este método, consideren leer el documento "Enunciado del análisis de Kleene" disponible en la página de EVA del curso.

[1] Esto es debido a que X_0 representa a las tiras que al ser reconocidas desde el estado inicial terminan en algún estado final.

1.2 Lema de Arden

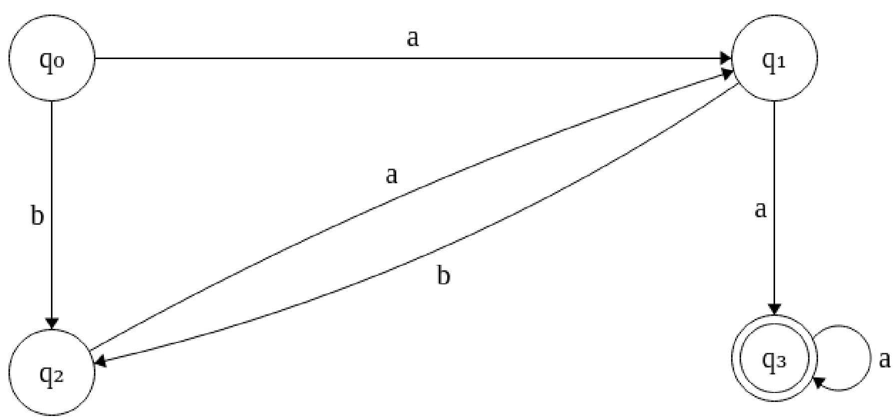
El Lema de Arden nos permite afirmar que $X = r^*.s$ es la solución para la ecuación

$$X = rX|s, \text{ donde } \epsilon \notin L(r)$$

Nuevamente, por más detalle, pueden leer la prueba formal de este resultado en la página de EVA que, a su vez, es la solución para el ejercicio 9 del Práctico 1.

2 El ejercicio

Considere el autómata finito determinista $M = (\{q_0, q_1, q_2, q_3\}, \Sigma = \{a, b\}, \delta, q_0, F = \{q_3\})$, ilustrado por el siguiente diagrama:



Encuentre una expresión regular r tal que $L(r) = L(M)$ utilizando el *Análisis de Kleene*.

3 La solución

Comenzamos planteando las ecuaciones X_i para cada estado q_i, siguiendo con atención sus **transiciones salientes**.

$X_0 = aX_1|bX_2$
$X_1 = aX_3|bX_2$
$X_2 = aX_1$
$X_3 = aX_3$

Sin embargo a X_3 le está faltando algo. Acorde a lo visto en la introducción, para X_3 la ecuación se forma a partir de la pregunta:

¿Cuáles son las tiras que, al ser reconocidas desde q_3, terminan en *algún* estado final?

Como q_3 es un estado final, una de la formas posibles para esas tiras es la tira vacía[2], así que la ecuación nos queda $X_3 = aX_3|\epsilon$[3].

Ahora, con el sistema de ecuaciones totalmente planteado

$X_0 = aX_1|bX_2$
$X_1 = aX_3|bX_2$
$X_2 = aX_1$
$X_3 = aX_3|\epsilon$

hallaremos las expresiones regulares X_0, X_1, X_2 y X_3 utilizando el **Lema de Arden** e identidades de expresiones regulares.

$X_3 = aX_3|\epsilon \overset{Lema\ de\ Arden}{\Longrightarrow} X_3 = a^*.\epsilon = a^*$

$X_1 = aX_3|bX_2 \overset{Sust.\ X_3}{\Longrightarrow} X_1 = a.a^*|bX_2 \overset{Sust.\ X_2}{\Longrightarrow} X_1 = a.a^*|b.aX_1$
$\overset{Conm.\ union}{=} baX_1|aa^* \overset{Lema\ de\ Arden}{\Longrightarrow} X_1 = (ba)^*.aa^*$

$X_2 = aX_1 \overset{Sust.\ X_1}{\Longrightarrow} X_2 = a.(ba)^*aa^*$

$X_0 = aX_1|bX_2 \overset{Sust.}{\Longrightarrow} X_0 = a.(ba)^*aa^*|b.a(ba)^*aa^*$

Por lo tanto, como X_0 representa al estado inicial, la expresión regular que buscamos es $r = X_0$ donde:

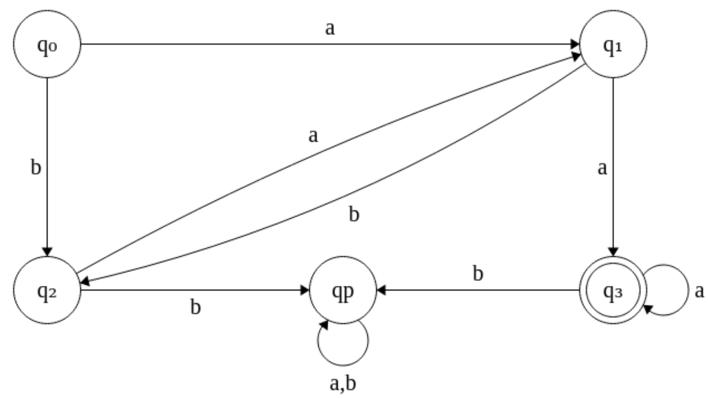
$$L(r) = L(X_0) = L(M) \Longleftrightarrow L(a(ba)^*aa^*|ba(ba)^*aa^*) = L(M)$$

[2] O, lo que es equivalente, que el autómata esté en el estado final y no consuma símbolos de la entrada.

[3] Como regla general, a cada ecuación que representa a un estado final, le agregamos la opción de la tira vacía. Por ejemplo, si en el caso de este ejercicio q_1 hubiese sido también final, la ecuación X_1 tendría la forma $X_1 = ...|\epsilon$.

4 ¿Y el pozo?

El autómata finito determinista M presentado en el ejercicio no tenía la función de transición definida en su totalidad. Para completarla deberíamos insertar un nuevo estado que, por su semántica, podríamos llamarlo q_p (estado pozo).



Entonces ahora nuestro sistema de ecuaciones sería muy similar al anterior pero agregando aquellas **transiciones salientes** hacia el estado pozo

$X_0 = aX_1|bX_2$
$X_1 = aX_3|bX_2$
$X_2 = aX_1|bX_p$
$X_3 = aX_3|bX_p|\epsilon$
$X_p = aX_p|bX_p$

Pero como desde el estado pozo no existe ningún patrón para tiras que puedan llegar a un estado final, la solución a X_p es $X_p = \emptyset$. Es decir, la expresión regular que genera al lenguaje vacío. Y además, como para toda expresión regular r se cumple que $r.\emptyset = \emptyset$, el sistema de ecuaciones vuelve a quedar igual a como estaba antes de agregarle el estado q_p al autómata

$X_0 = aX_1|bX_2$
$X_1 = aX_3|bX_2$
$X_2 = aX_1|bX_p \overset{Sust.\ X_p}{=} aX_1|b.\emptyset = aX_1|\emptyset = aX_1$
$X_3 = aX_3|bX_p|\epsilon \overset{Sust.\ X_p}{=} aX_3|b.\emptyset|\epsilon = aX_3|\emptyset|\epsilon = aX_3|\epsilon$
$X_p = \emptyset$

Por lo tanto, para el **Análisis de Kleene, no es necesario incluir al pozo en el sistema de ecuaciones**, porque la solución para X_p es siempre la misma:

$$X_p = \emptyset$$