Theory of automata and Formal languages

Fernando Javier López Cerezo
Taller 1

October 14, 2022

Una expresión regular es ambigua si una cadena puede ser generada de dos formas diferentes a partir de la misma expresión regular. Especifica si las siguientes expresiones regulares sobre el alfabeto $\{a, b, c, d\}$ son ambiguas o no, y explica claramente el por qué (razona qué lenguaje generan y demuestra o refuta adecuadamente tu respuesta):

a)

Sea $R = a((ab)^*cd)^* + a(ababcb^*)a^*$. $L(R) = L(R_1) \cup L(R_2)$ donde $R_1 = a((ab)^*cd)^*$ y $R_2 = a(ababcb^*)a^*$. Estudiemos estos dos lenguajes por separado:

$L(R_1) = L(a((ab)^*cd)^*) = L(a)L((ab)^*cd)^* = L(a)(L(ab)^*L(cd))^* =$
$= \{a, acd, acdcd, ..., aabcd, aabcdcd, ..., aababcd, aababcdcd, ...\}$

$L(R_2) = L(a(ababcb^*)a^*) = L(a)L(ababcb^*)L(a^*) = L(a)L(ababc)L(b)^*L(a)^* =$
$= \{aababc, aababca, aababcba, aababcbaa, aababcbaaa, ..., aababcbb, aababcbba, aababcbbaa, aababcbbaaa, ...\}$

Veáse que $L(R_1) \cap L(R_2) = \emptyset$ pues en toda cadena de $L(R_2)$ aparece el símbolo c (menos en a pero a $\notin L(R_2)$) y en toda cadena de $L(R_1)$ cuando aparece el símbolo c viene acompañado del símbolo d (cd) cosa que en $L(R_2)$ no es posible ya que ni siquiera usa el símbolo d.

Luego en $L(R)$ toda cadena es generada por $L(R_1)$ ó $L(R_2)$, al ser su intersección vacía concluimos que no hay cadena que se pueda generar de dos formas distintas, luego R NO es ambigua.

b)

Sea $R = aab^*(ab)^* + ab^* + a^*bba^*$. $L(R) = L(R_1) \cup L(R_2) \cup L(R_3)$ donde $R_1 = aab^*(ab)^*$, $R_2 = ab^*$ y $R_3 = a^*bba^*$. En esta caso basta con estudiar $L(R_2)$ y $L(R_3)$.

$L(R_2) = L(a)L(b)^* = \{a, ab, abb, abbb, ...\}$
$L(R_3) = L(a)^*L(bb)L(a)^* = \{bb, abb, bba, abba, aabb, ...\}$

Veáse que la cadena abb pertenece tanto a $L(R_2)$ como a $L(R_3)$ luego R es una expresión regular ambigua ya que la cadena abb puede ser generada de dos formas diferentes (utilizando R_2 ó R_3).

c)

Sea $R = aaba^* + aaaba + aabba^* + a$. $L(R) = L(R_1) \cup L(R_2) \cup L(R_3) \cup L(R_4)$ donde $R_1 = aaba^*$, $R_2 = aaaba$, $R_3 = aabba^*$ y $R_4 = a$.

$L(R_1) = L(aab)L(a)^* = \{aab, aaba, aabaa, ...\}$
$L(R_2) = L(aaaba) = \{aaaba\}$
$L(R_3) = L(aabb)L(a)^* = \{aabb, aabba, aabbaa, ...\}$
$L(R_4) = L(a) = \{a\}$

Es claro que estos cuatro lenguajes son disjuntos dos a dos luego R es una expresión regular NO ambigua.

Nota)

En los ejercicios anteriores sólo se ha comprobado que si el lenguaje $L(R)$ formado por una expresión regular R es en realidad unión de varios lenguajes $L(R_i)$ que su intersección sea vacía para ver que no hay varias formas de generar una misma cadena utilizando los distintos lenguajes. Véase que en realidad una expresión regular como a^*a^* podría generar la misma cadena de formas diferentes (como por ejemplo $a\epsilon$ y ϵa) luego habría que comprobar también que este no sea el caso en cada uno de los $L(R_i)$. Para estos ejemplos considero que es claro que no es el caso y que se puede ver a ojo pero considero necesario comentarlo para dar el ejercicio por completo.