

# Teoría de Lenguajes - Lema de Arden

## 1. Enunciado

Sean  $r$  y  $s$  expresiones regulares. Considere la ecuación  $X = r.X|s$ , donde  $r.X$  denota la concatenación de  $r$  con  $X$ , y  $|$  la unión. Asumiendo que el conjunto denotado por  $r$  no contiene la tira vacía, encuentre la solución para  $X$  y pruebe que es única. ¿Cuál es la solución si  $L(r)$  contiene la tira vacía?

## 2. Solución

Intuitivamente, podemos ver que todas las tiras de  $L(s)$  pertenecen a la solución, así como las tiras de  $L(rs)$ , y así hasta  $L(r^*s)$ , que sería el “límite”.

Vemos primero que  $r^*s$  es solución de la ecuación, es decir que  $L(r^*s) = L(rr^*s|s)$ :

$$L(r(r^*s)|s) = L(rr^*s|s) = L((rr^*|\epsilon)s) = L(r^*s)$$

(Todas justificadas por propiedades de expresiones regulares fácilmente demostrables)

Demostremos ahora que la solución es única. Primero, vemos que cualquier  $X'$  que sea solución debe cumplir que  $L(X')$  incluye a  $L(r^*s)$ . Mostraremos que cualquier tira de la forma  $w_{11} \dots w_{1k}w_2$ , con  $k \geq 0$ ,  $w_{1i} \in L(r)$ ,  $w_2 \in L(s)$  (es decir, las tiras de  $L(r^*s)$ ), pertenece a  $L(X')$ . Lo hacemos por inducción completa en  $k$ :

Paso Base : Si  $k = 0$ , entonces  $w \in L(s)$  y por lo tanto  $w \in L(X')$

Hipótesis Inductiva : Si  $k < i$ , entonces  $w \in L(X')$

Tesis Inductiva : Si  $k = i$ , entonces  $w \in L(X')$ . Esto es fácil de ver:  $w = w_{11}w'$ , con  $w_{11} \in L(r)$ , y  $w' \in L(r^*s)$ ; por lo tanto  $w'$  es más corta que  $w$  (porque  $\epsilon \notin L(r)$ ), y por HI  $w' \in L(X')$ . Como  $X'$  cumple la ecuación, se cumple que  $w \in L(X')$

Ahora, supongamos que  $L(X')$  incluye estrictamente a  $L(r^*s)$ . Sea  $w \in L(X')$  tal que  $w \notin L(r^*s)$ , y no existe una tira más corta que  $w$  que cumpla esta condición. Como  $w \in L(X')$ , entonces existen dos posibilidades:

- Si  $w \in L(s)$ , entonces  $w \in L(r^*s)$ , lo que es absurdo

- Si  $w \in L(rX')$ , entonces  $w = w_1w_2$ , con  $w_1 \in L(r)$  y  $w_2 \in L(X')$ ; al ser  $w_2$  más corta que  $w$  (porque  $\epsilon \notin r$ ),  $w_2 \in L(r^*s)$ ; pero entonces  $w \in L(r^*s)$ , lo cual es absurdo.

Todas las opciones conducen a un absurdo que se origina al suponer que existe una solución  $X'$  que incluye estrictamente a  $r^*s$ , luego  $r^*s$  es la única solución.

Finalmente, vemos qué sucede si  $\epsilon \in L(r)$ . En este caso, la ecuación se transforma en:  $X = r'X|s|X$  donde  $L(r') = L(r) - \epsilon$ . La solución para esta ecuación es  $r^*(\alpha|s)$  para cualquier e.r.  $\alpha$ . Por lo tanto, la solución ya no es única (salvo que  $L(s) = \Sigma^*$ ).