Autómatas y Lenguajes

3er curso

1er cuatrimestre

Alfonso Ortega: alfonso.ortega@uam.es

UNIDAD 1: Modelos de cómputo y familias de lenguajes

TEMA 3: Autómatas finitos deterministas y no deterministas, expresiones regulares

Expresiones regulares

Expresiones regulares

Introducción

- Kleene introdujo en 1956 [Kleene56] el concepto de expresión regular para expresar el conjunto de palabras aceptadas por los autómatas finitos.
- Dado un alfabeto Σ se representará mediante E_{Σ} el conjunto de todas las expresiones regulares definidas sobre Σ.

Expresiones regulares

Definición

- Dados
  - Un alfabeto Σ
  - La siguiente notación (previamente definida)
    - Conjunto vacío: Φ
    - Palabra vacía: λ
    - Unión de lenguajes: +
    - Concatenación de lenguajes y palabras: .
    - Cierre o clausura de lenguajes: *
- Se define
  1. $\Phi \in E_\Sigma$
  2. $\lambda \in E_\Sigma$
  3. $\forall a \in \Sigma,\ a \in E_\Sigma$
  4. $\alpha+\beta,\ \alpha.\beta \in E_\Sigma \quad \forall \alpha, \beta \in E_\Sigma$
  5. $\alpha^* \in E_\Sigma \quad \forall \alpha \in E_\Sigma$
  6. Sólo son expresiones regulares las que se obtienen por aplicación, un número finito de veces, de las reglas anteriores.

Expresiones regulares

Ambigüedad: prioridad entre los operadores

- El orden de prioridad de los operadores es el siguiente:
  - Cierre o clausura de lenguajes: *
  - Concatenación de lenguajes y palabras: .
  - Unión de lenguajes: +
- El orden puede modificarse con el uso de paréntesis con la notación matemática habitual.

Expresiones regulares

Ejemplos

- La tabla siguiente muestra algunos ejemplos, la primera columna contiene los alfabetos y la segunda las expresiones regulares.

Σ	**Expresión regular**
$\Sigma=\{a,b,c,...,z\}$	$(a+b+c+...+z)^*$
$\Sigma=\{0,1\}$	0^*10^*
	$01+000$
	$(0+1)^*00\ (0+1)^*$
$\Sigma=\{a,b,c\}$	$a(a+b+c)^*$
	$a+bc+b^2a$
$\Sigma=\{a,b\}$	$a^*(a+b)$
	$(a+b)^*(a+bb)$
	$(aa)^*(bb)^*b$

Lenguaje representado por una expresión regular

Definición

- Dados
  - Un alfabeto Σ
  - Se define **lenguaje de una expresión regular** r ($\forall r \in E_\Sigma$) y se escribe $L(r)$ ($L(r) \subseteq \Sigma^*$)
    - $L(\Phi)=\Phi$
    - $L(\lambda)=\{\lambda\}$
    - $\forall a\in\Sigma,\ L(a)=\{a\}$
    - $\forall\alpha,\ \beta \in E_\Sigma$
      - $L(\alpha+\beta)=L(\alpha)\cup L(\beta)$
      - $L(\alpha.\beta)=L(\alpha).L(\beta)$
    - $\forall\alpha \in E_\Sigma,\ L(\alpha^*)=L(\alpha)^*$.

- A continuación se verá que E_Σ también representa a todos los lenguajes regulares.

Lenguaje representado por una expresión regular

Ejemplos

- La tabla siguiente muestra algunos ejemplos, la primera columna contiene los alfabetos, la segunda las expresiones regulares y la tercera los lenguajes.

Σ	**Expresión regular**	**Lenguaje**
$\Sigma=\{a,b,c,...,z\}$	$(a+b+c+...+z)^*$	Σ^*
$\Sigma=\{0,1\}$	0^*10^*	Cadenas binarias que contienen un solo 1
	$01+000$	$\{01, 000\}$
	$(0+1)^*00\ (0+1)^*$	Cadenas binarias que contienen al menos un par de 0 consecutivos
$\Sigma=\{a,b,c\}$	$a(a+b+c)^*$	Cadenas formadas por a, b o c que comienzan por a
	$a+bc+b^2a$	$\{a,bc,b^2a\}$
$\Sigma=\{a,b\}$	$a^*(a+b)$	$\{a,aa, aaa, a...a, b, ab, aab, aaab, a...ab\}$
	$(a+b)^*(a+bb)$	Cadenas formadas por a o b que terminan en a o en bb
	$(aa)^*(bb)^*b$	Cadenas formadas por un número par de símbolos a (incluso 0) seguidos por un número impar de símbolos b (al menos 1) = $\{a^{2n}b^{2m+1}: n\geq 0, m\geq 0\}$

Lenguaje representado por una expresión regular

Observaciones prácticas

- Por la definición anterior podemos observar lo siguiente.
  - Supóngase

$$\Sigma=\{a, b\}$$

  - Entonces, si se reflexiona respecto a $(a+b)^*$.

$$L((a+b)^*)=(L(a+b))^*$$

  - Pero por definición de + y de Σ

$$L(a+b)=\Sigma$$

  - Por lo que

$$L((a+b)^*)=\Sigma^*$$

  - Ahora bien, Σ^* es el conjunto de todas las cadenas que se pueden formar con símbolos de Σ y, por lo tanto, el mayor lenguaje posible con símbolos de ese alfabeto.
- En general

$$\Sigma=\{a_1, a_2,..., a_n\} \Rightarrow \Sigma^*=L((a_1+a_2+...+a_n)^*)$$

- Y la expresión regular “más grande” (en el sentido de que represente al lenguaje mayor) que se puede definir sobre Σ es $(a_1+a_2+...+a_n)^*$

Equivalencia de expresiones regulares: definición

- Dos expresiones regulares α y β son equivalentes y se representa

$$\alpha=\beta$$

- si y sólo si describen el mismo lenguaje, es decir,

$$L(\alpha)= L(\beta)$$

Equivalencia de expresiones regulares: propiedades

- Se pueden demostrar, algunas son evidentes a partir de la definición de las expresiones regulares, las siguientes propiedades que sirven para modificarlas algebraicamente sin cambiar el lenguaje que representan, es decir, manteniendo la equivalencia entre ellas.
  1. + es asociativa: $\alpha+(\beta+\delta)=(\alpha+\beta)+\delta \ \forall\alpha, \beta, \delta \in E_\Sigma$
  2. + es conmutativa: $\alpha+\beta=\beta+\alpha \ \ \forall\alpha, \beta \in E_\Sigma$
  3. . es asociativa: $\alpha.(\beta.\delta)=(\alpha.\beta).\delta \ \forall\alpha, \beta, \delta \in E_\Sigma$
  4. . es distributiva respecto + : $\forall\alpha, \beta, \delta \in E_\Sigma$

     $$\alpha.(\beta+\delta)=(\alpha.\beta)+(\alpha.\delta)$$

     $$(\beta+\delta).\alpha =(\beta.\alpha)+(\delta.\alpha)$$
  5. Existencia del elemento neutro de . (λ) : $\alpha.\lambda=\lambda.\alpha=\alpha \ \forall\alpha \in E_\Sigma$
  6. Existencia del elemento neutro de + (Φ) : $\alpha+\Phi = \Phi+\alpha=\alpha \ \forall\alpha \in E_\Sigma$
  7. Cierre de λ, aplicando la propiedad 5 y ya que $L(\lambda)=\{\lambda\}$: $\lambda^*=\lambda$
  8. Concatenación con Φ, por la definicion de . en lenguajes y ya que $L(\Phi)=\Phi$:

     $$\alpha.\Phi = \Phi.\alpha=\Phi \ \forall\alpha \in E_\Sigma$$

Expresiones regulares

Equivalencia de expresiones regulares: propiedades

9. Cierre de Φ, por la definición de cierre: $\Phi^*=\lambda$
10. Cierre y concatenación: $\alpha^*.\alpha^*=\alpha^* \; \forall\alpha \in E_\Sigma$
11. Cierre y concatenación: $\alpha.\alpha^*=\alpha^*.\alpha \; \forall\alpha \in E_\Sigma$
12. Cierre del cierre: $(\alpha^*)^*=\alpha^* \; \forall\alpha \in E_\Sigma$
13. Expresión del cierre como uniones:

$$\alpha^* = \bigcup_{i=0}^{n} \alpha^i + \alpha^{n+1}.\alpha^* \quad \forall\alpha \in E_\Sigma, n \in N$$

$$\alpha^* = \lambda + \bigcup_{i=1}^{n} \alpha^i + \alpha^{n+1}.\alpha^*$$

14. Expresión (mínima) del cierre como uniones: $\alpha^*=\lambda+\alpha.\alpha^* \; \forall\alpha \in E_\Sigma$
15. Expresión del cierre como uniones, aplicando reiteradas veces la propiedad anterior y reordenando: $\alpha^*=(\lambda+\alpha)^{n-1}+ \alpha^n.\alpha^* \; \forall\alpha \in E_\Sigma, n \in N$
16. Cierre de la unión: (se puede entender como que nada -ninguna expresión regular- puede añadir nada a la "más grande")

$$f(\alpha_1, \alpha_2,...,\alpha_n)+(\alpha_1+\alpha_2+...+\alpha_n)^*=(\alpha_1+\alpha_2+...+\alpha_n)^* \; \forall\alpha_i \in E_\Sigma, \; \forall f{:}E_\Sigma^{\,n}\rightarrow E_\Sigma$$

Expresiones regulares

Equivalencia de expresiones regulares: propiedades

17. Cierre de expresiones del cierre de cualquier expresión regular:

$$(f(\alpha_1^*, \alpha_2^*,...,\alpha_n^*))^*=(\alpha_1+\alpha_2+...+\alpha_n)^* \; \forall\alpha_i \in E_\Sigma, \; \forall f{:}E_\Sigma^{\,n}\rightarrow E_\Sigma$$

- Donde se debe entender que *f* es una función que no puede expresarse con menos variables de las que aparecen
- Se puede entender que el cierre de cualquier expresión regular en la que todas las variables que incluya aparecen cerradas es la expresión regular "más grande" que se puede describir con las expresiones regulares de partida.

18. Como consecuencia de la anterior: $(\alpha^*+\beta^*)^*=(\alpha^*.\beta^*)^*=(\alpha+\beta)^* \;\; \forall\alpha, \beta \in E_\Sigma$
19. Cierre y concatenación, generalización: $(\alpha.\beta)^*.\alpha = \alpha.(\beta.\alpha)^* \; \forall\alpha, \beta \in E_\Sigma$
20. Cierre y concatenación: $(\alpha^*.\beta)^*.\alpha^* = (\alpha+\beta)^* \; \forall\alpha, \beta \in E_\Sigma$
21. Cierre y concatenación: $(\alpha^*.\beta)^* = (\alpha+\beta)^*.\beta+\lambda \; \forall\alpha, \beta \in E_\Sigma$
22. Reglas de inferencia (útiles para el problema de análisis por Arden):

$$r=s^*.t \Rightarrow r=s.r+t \; \forall r, s, t \in E_\Sigma$$

$$\lambda\notin L(s) \; r=s.r+t \; \Rightarrow r=s^*.t \; \forall r, s, t \in E_\Sigma$$