Expresiones Regulares y Gramáticas Regulares

Sistemas Lineales.

Universidad de Cantabria

Esquema

Problema

Nos preguntamos si las expresiones regulares generan los mismos lenguajes que las gramáticas regulares.

Sabemos que a partir de una expresión regular podemos hallar una gramática regular que genera el mismo lenguaje, pero no sabemos el recíproco.

Problema

Nos preguntamos si las expresiones regulares generan los mismos lenguajes que las gramáticas regulares.
Sabemos que a partir de una expresión regular podemos hallar una gramática regular que genera el mismo lenguaje, pero no sabemos el recíproco.

Ejemplo Sencillo

Si tenemos la siguiente gramática regular:

$$
\begin{aligned}
q_0 &\mapsto aq_1, \\
q_1 &\mapsto aq_1 | b.
\end{aligned}
$$

¿Existe una expresión regular que lo genera?

Cosas que podemos deducir: las palabras empiezan por “a”, terminan en “b”...

Ejemplo Sencillo

Si tenemos la siguiente gramática regular:

$$\begin{aligned} q_0 &\mapsto aq_1, \\ q_1 &\mapsto aq_1|b. \end{aligned}$$

¿Existe una expresión regular que lo genera?
Cosas que podemos deducir: las palabras empiezan por “a”, terminan en “b”...

Ejemplo Sencillo

Cosas que podemos hacer:
Consideremos diferentes gramáticas, donde las producciones son las mismas pero tienen diferentes símbolos iniciales y a cada uno de ellos se le asocia un lenguaje.

Ejemplo Sencillo

Obtengo esta ecuación de conjuntos:

$$\begin{aligned} L_0 &= aL_1, \\ L_1 &= aL_1 \cup \{b\}. \end{aligned}$$

Ejemplo Sencillo

Supongamos que estos lenguajes están generados por lenguajes regulares:

$$\begin{aligned} X_0 &= aX_1, \\ X_1 &= aX_1 + b. \end{aligned}$$

Vamos a suponer que todos estos lenguajes son dados por una expresión regular, ¿como hallar una expresión regular que lo genere?

Ejemplo Sencillo

Supongamos que estos lenguajes están generados por lenguajes regulares:

$$\begin{aligned} X_0 &= aX_1, \\ X_1 &= aX_1 + b. \end{aligned}$$

Vamos a suponer que todos estos lenguajes son dados por una expresión regular, ¿como hallar una expresión regular que lo genere?

Ecuaciones Lineales

Definición

Llamaremos sistema de ecuaciones lineales en expresiones regulares a toda ecuación del tipo siguiente:

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},$$

donde los $\alpha_{i,j}$ y los β_k son expresiones regulares sobre un alfabeto Σ.

Ecuación Fundamental

Definición

Se denomina ecuación lineal fundamental en expresiones regulares a la ecuación lineal en una variable X siguiente:

$$X = \alpha X + \beta,$$

donde α y β son expresiones regulares sobre un alfabeto finito Σ.

Lema de Arden

Teorema (Lema de Arden)

Dada la ecuación fundamental siguiente:

$$X = \alpha X + \beta,$$

donde α, β son expresiones regulares sobre un alfabeto Σ. Se verifican las propiedades siguiente:

1. *La ecuación fundamental anterior posee una solución única si y solamente si $\lambda \notin L(\alpha)$.*
2. *La expresión regular $\alpha^* \cdot \beta$ es siempre solución de la ecuación fundamental anterior.*
3. *Si $\lambda \in L(\alpha)$, para cualquier expresión regular γ, la expresión $\alpha^* \cdot (\beta + \gamma)$ es una solución de la ecuación fundamental.*

Resolución Gausiana

El caso $n = 1$ se resuelve mediante el Lema de Arden. Para el caso $n > 1$, usaremos un doble paso:

Despejar. Podemos despejar X_n en la última ecuación, mediante la expresión siguiente:

$$X_n := \alpha_{n,n}^{*} R_n,$$

donde $R_n := \sum_{j=1}^{n-1} \alpha_{n,j} X_j + \beta_n$.

Resolución Gausiana

Sustituir. Podemos sustituir la expresión anterior en el resto de las ecuaciones obteniendo un nuevo sistema de $(n-1)$ ecuaciones en $(n-1)$ variables. Este sistema viene dado, obviamente, por las expresiones siguientes para $1 \leq i \leq n-1$:

$$X_i := \left(\sum_{j=1}^{n-1} \left(\alpha_{i,j} + \alpha_{i,n}\alpha_{n,n}^{*}\alpha_{n,j} \right) X_j \right) + \left(\beta_i + \alpha_{i,n}\alpha_{n,n}^{*}\beta_n \right).$$

Resolución Gausiana

Levantamiento. Una vez llegados al caso $n = 1$, se obtiene una expresión regular válida para X_1 y se procede a levantar el resto de las variables usando las expresiones obtenidas en la fase de despejado.

Sistema Lineal asociado a una Gramática

Supongamos $V = \{q_0, \ldots, q_n\}$ es el conjunto de los símbolos no terminales, que supondremos de cardinal $n + 1$. Definamos un conjunto de variables $\{X_0, \ldots, X_n\}$ con el mismo cardinal y con la asignación $q_i \mapsto X_i$ como biyección.

Sistema Lineal asociado a una Gramática

Definamos para cada i, $0 \leq i \leq n$, la expresión regular β_i mediante la construcción siguiente. Consideremos todas las producciones que comienzan en la variable q_i y terminan en elementos de $\Sigma \cup \{\lambda\}$. Supongamos que tales producciones sean

$$q_i \mapsto a_1 \mid \ldots \mid a_r.$$

Definimos

$$\beta_i := a_1 + \ldots + a_r.$$

Si no hubiera ninguna producción del tipo $q_i \mapsto a \in \Sigma \cup \{\lambda\}$, definiremos $\beta_i := \emptyset$.

Sistema Lineal asociado a una Gramática

Para cada i y para cada j, definiremos el coeficiente $\alpha_{i,j}$ del modo siguiente. Consideremos todas las producciones que comienzan en el símbolo no terminal q_i e involucran al símbolo no terminal q_j. Supongamos que tales producciones sean:

$$q_i \mapsto a_1 q_j \mid \cdots \mid a_r q_j,$$

con $a_k \in \Sigma \cup \{\lambda\}$. Entonces definiremos

$$\alpha_{i,j} := a_1 + \cdots + a_r.$$

Si no hubiera ninguna de tales producciones, definiremos $\alpha_{i,j} := \emptyset$.

Sistema Lineal asociado a una Gramática

Definición (Sistema asociado a una gramática)

Dada una gramática $G = (V, \Sigma, q, P)$ llamaremos sistema asociado a G y lo denotaremos por $\mathcal{S}(G)$ al sistema:

$$\mathcal{S}(G) := \left\{ \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \alpha_{0,1} & \cdots & \alpha_{0,n} \\ \vdots & \ddots & \vdots \\ \alpha_{n,0} & \cdots & \alpha_{n,n} \end{pmatrix} \begin{pmatrix} X_0 \\ \vdots \\ X_n \end{pmatrix} + \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} \right\},$$

dado por las anteriores reglas de construcción.

Equivalencia entre ER y GR

Teorema

Con las anteriores notaciones, sea $(\alpha_0, \ldots, \alpha_n)$ *una solución del sistema* $S(G)$ *asociado a una gramática* G*. Entonces,* $L(\alpha_0)$ *es el lenguaje generado por la gramática* G*.*