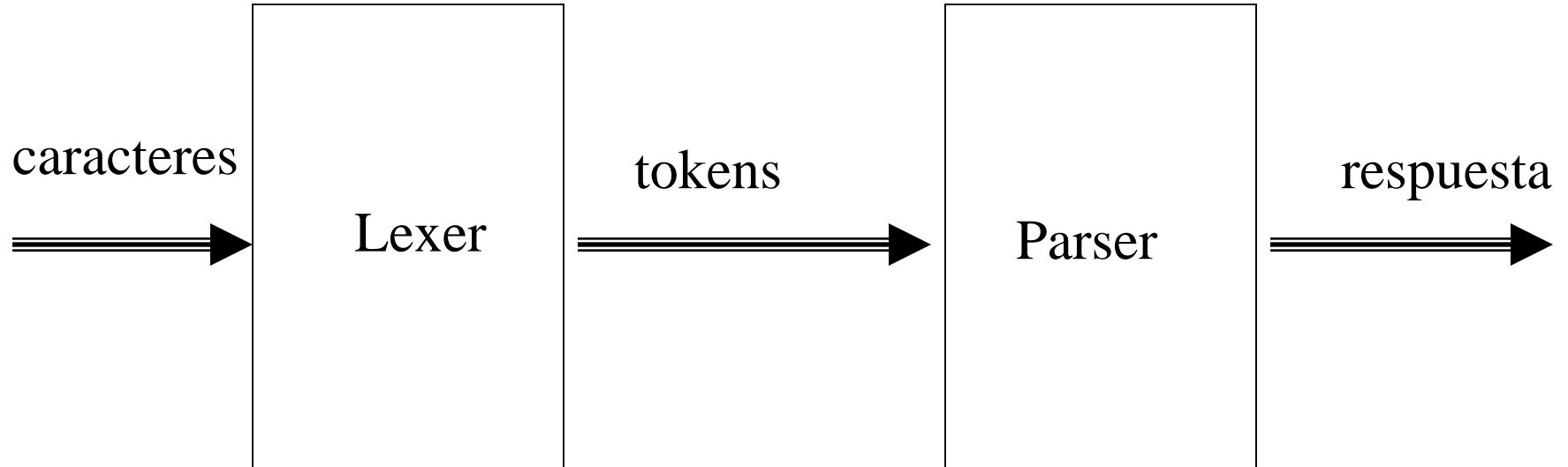


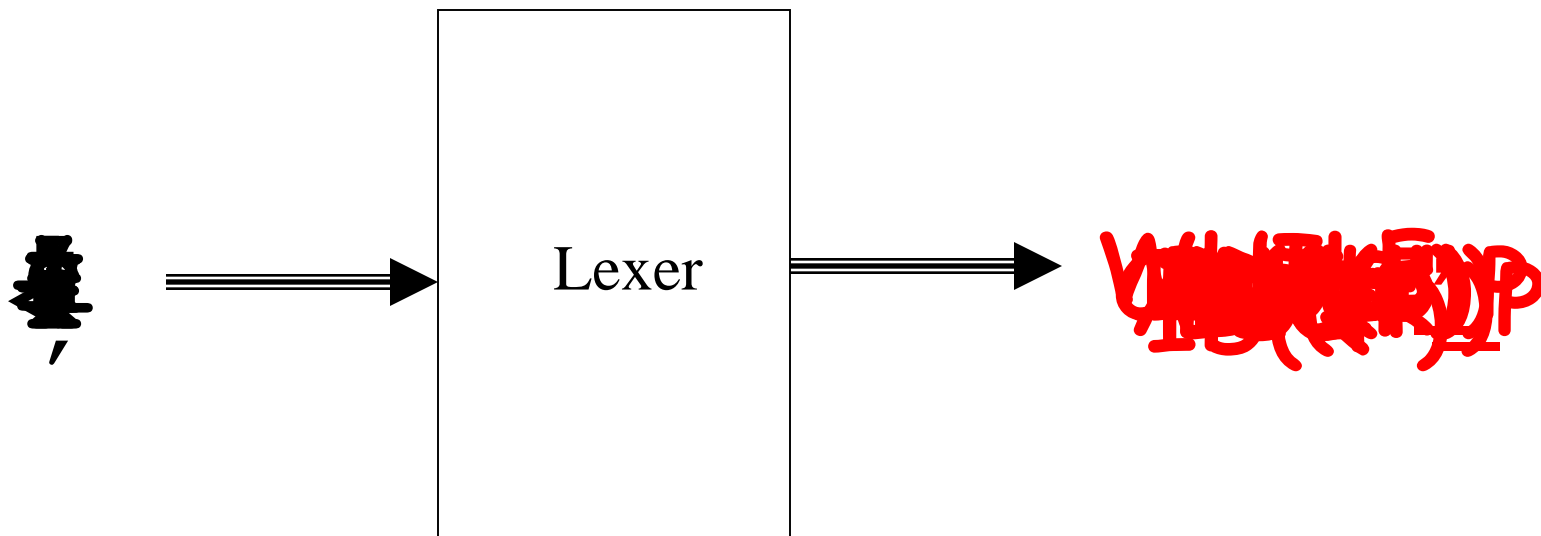
Sintaxis y Semántica



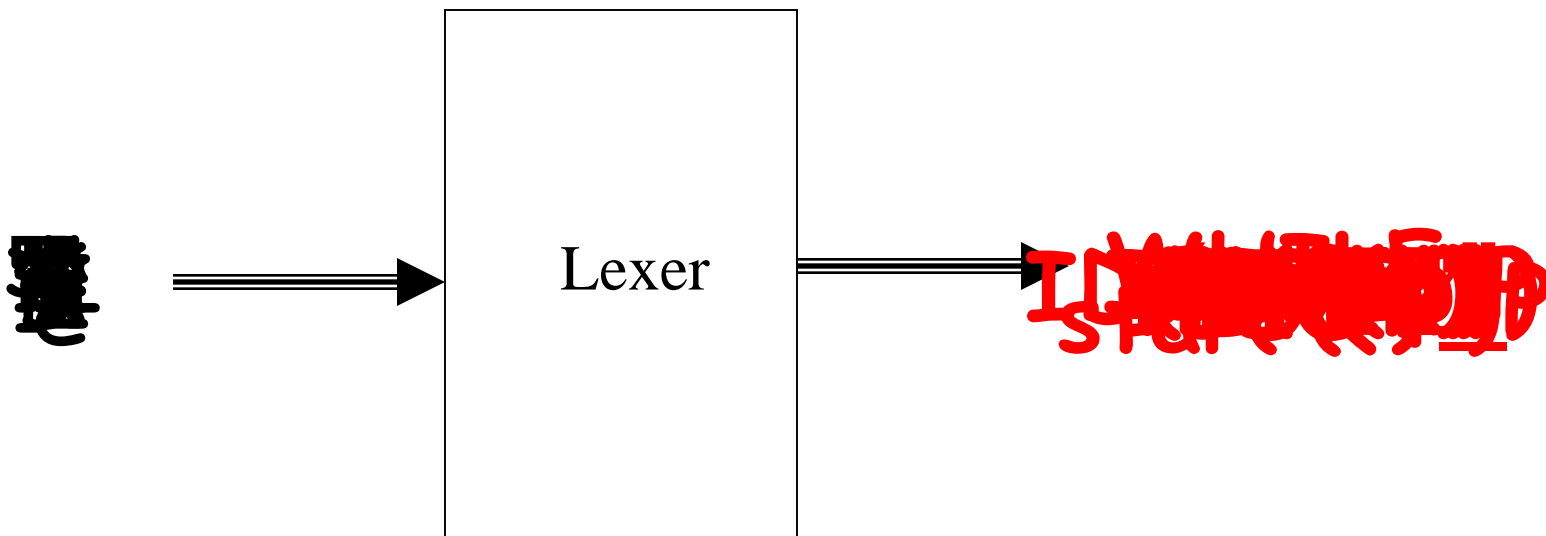
Fases en el proceso de análisis de lenguajes



```
while (i<=5) x = x+i ;
```



```
while (i<5) x = x+i 55 temp *
```



Definición: Alfabeto

Un conjunto finito de símbolos:

- {a,b,c,d}
- {0,1,2,3,4,5,6,7,8,9}
- {identificador, número, +, -, *, /}
- {while, if, '{', '}', >, >=, <, <=}
- { Tipo Transacción, IdSucursal, NumCuenta, Cantidad}

Definición: cadena o palabra de un alfabeto

Secuencia finita de símbolos de un alfabeto

Definición recursiva (Cadena de un alfabeto A)

- *La cadena vacía (λ) es una cadena sobre cualquier alfabeto*
- *Si β es una cadena de A y $\sigma \in A$ entonces $\sigma\beta$ es una cadena de A*

Ejemplo: cadena o palabra de un alfabeto

$A = \{id, num, +, -, *, /, (,)\}$

$id * (id + id * num - num)$

$id * (id + id * num id id num()$

\wedge

Ejemplo: cadena o palabra de un alfabeto

$A = \{a, b, c\}$

a

aab

acab

\wedge

Operaciones sobre cadenas:

Longitud: $\#$

$$\#\lambda = 0$$

$$\#\sigma\beta = 1 + \#\beta$$

Concatenación: $\omega\beta$

$$\lambda = 0$$

$$(\sigma\beta)\omega = \sigma(\beta\omega)$$

Definición: Lenguaje



Conjunto de cadenas de un alfabeto.

Sintaxis



- Forma correcta de escribir programas
- Formalismo para describir la forma de los programas bien escritos
- Gramáticas - reglas que describen cómo construir todas las palabras en un lenguaje dado
- Ejemplos

Semántica



- Significado de las estructuras semánticas del programa
- Describe el comportamiento de los programas
- Describe cómo debe traducirse al lenguaje de máquina

Gramáticas: formalismo para definir sintaxis

Define la forma que deben tener los las cadenas que pertenecen a un lenguaje

Generalmente se usan para determinar si un programa es correcto sintácticamente

También pueden usarse para la reingeniería reversa y para la reestructuración

Un ejemplo más fácil

■ Una expresión aritmética entera

- Un número o una constante entera
- Si A y B son expresiones aritméticas las siguientes también son expresiones aritméticas

$$\Rightarrow A + B$$

$$\Rightarrow A * B$$

$$\Rightarrow A - B$$

$$\Rightarrow A / B$$

$$\Rightarrow A \% B$$

$$\Rightarrow (A)$$

$$\Rightarrow -A$$

Definición 1: Gramática Independiente del contexto

Una gramática es una 4-tupla (N, T, S, P) donde:

*N es un conjunto de símbolos **no-terminales**,*

*T es un conjunto de símbolos **terminales**,*

*S es un símbolo no-terminal denominado **símbolo distinguido** ($S \in N$)*

*P es un conjunto de **producciones**.*

Producciones

Una producción es una regla con de la forma

$\beta \rightarrow \delta$ donde:

- $A \in N$,
- $\delta \in (N \cup T)^*$
- $B \in (N \cup T)^*$.

Definición 3: Derivable en 1 paso

Dada una gramática independiente del contexto con $(\omega \rightarrow \varphi)$ una producción.

La producción indica que si tenemos una cadena de la forma $\beta\omega\delta$, podemos obtener $\beta\varphi\delta$, reemplazando ω por φ .

En este caso, se dice que $\beta\varphi\delta$ es derivable en un paso de $\beta\omega\delta$

y escribimos: $\beta\omega\delta \Rightarrow \beta\varphi\delta$.

Derivable

Dadas dos cadenas ω_0 y ω_n , decimos que ω_n es derivable (en cero o más pasos) de ω_0 , y escribimos $\omega_0 \Rightarrow^* \omega_n$ si se puede llegar de ω_0 a ω_n , por sucesivos reemplazos usando reglas de producción.

Lenguaje generado

Dada una gramática independiente del contexto $G=(N,T,S,P)$, el lenguaje generado por la gramática es el conjunto de todas las cadenas derivables a partir de S

$$L(G) = \{ \omega : \omega \in T^*, S \Rightarrow^* \omega \}$$

El ejemplo

■ $G = (\{E, T, F, S\}, \{\text{num}, \text{id}, +, (,), -\}, E, P)$

■ P:

- $E \rightarrow E + T$

- $E \rightarrow T$

- $T \rightarrow T * F$

- $T \rightarrow F$

- $F \rightarrow S$

- $F \rightarrow - S$

- $S \rightarrow (E)$

- $S \rightarrow \text{id}$

- $S \rightarrow \text{num}$

Ejemplo $35 + (10 + 30 * 4)$

E	$\Rightarrow E + T$	$\Rightarrow E + (E + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + F$	$\Rightarrow E + (T + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + (E)$	$\Rightarrow E + (F + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + (E + T)$	$\Rightarrow E + (\text{num}(10) + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + (E + T * F)$	$\Rightarrow T + (\text{num}(10) + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + (E + T * \text{num}(4))$	$\Rightarrow F + (\text{num}(10) + \text{num}(30) * \text{num}(4))$
	$\Rightarrow E + (E + F * \text{num}(4))$	$\Rightarrow \text{num}(35) + (\text{num}(10) + \text{num}(30) * \text{num}(4))$