

Sesión 16

Ambigüedad

Ambigüedad

- Si una gramática genera más de una estructura a partir de la misma raíz y con la misma cosecha (más de una estructura para la misma cadena), dicha gramática es ambigua
- Dos tipos de ambigüedad
 - En la gramática
 - En el lenguaje

Dr. Luis Pineda, IIMAS, 20102

Ambigüedad

- Si una gramática es ambigua, posiblemente (no necesariamente) existe una gramática no ambigua que genere el mismo lenguaje
- Un lenguaje es inherentemente ambiguo si todas sus gramáticas son ambiguas
- ¡No existe un algoritmo que decida si una gramática es ambigua!

Dr. Luis Pineda, IIMAS, 20103

Una gramática ambigua

- Exp es una GLC
 - $G_{exp} = (\{E\}, \{+, *, (,), 1, \dots, 9\}, E, P)$
donde $P = \{E \rightarrow E + E \mid E * E \mid (E) \mid 1 \mid \dots \mid 9\}$
- Una expresión ambigua:
 - $E + E * E$
- Dos derivaciones:
 - $E \Rightarrow E + E \Rightarrow E + E * E$
 - $E \Rightarrow E * E \Rightarrow E + E * E$
- ¡Son iguales!

Dr. Luis Pineda, IIMAS, 20104

Una gramática ambigua

- La expresión final es la misma:
 - $E \Rightarrow^* E + E * E$
 - $E \Rightarrow^* E + E * E$
- Pero las derivaciones son diferentes:
 - $E \Rightarrow E + E \Rightarrow E + E * E$
 - $E \Rightarrow E * E \Rightarrow E + E * E$

Dr. Luis Pineda, IIMAS, 20105

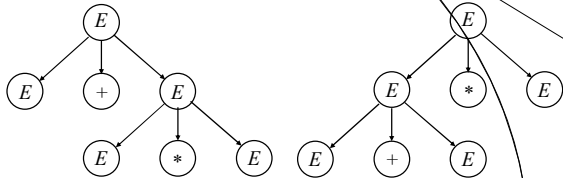
Una gramática ambigua

- A cada derivación le corresponde una estructura sintáctica:
 - Derivaciones diferentes pueden generar la misma estructura (para la misma cadena)
 - La ambigüedad surge cuando derivaciones diferentes generan estructuras diferentes (para la misma cadena)

Dr. Luis Pineda, IIMAS, 20106

Una gramática ambigua

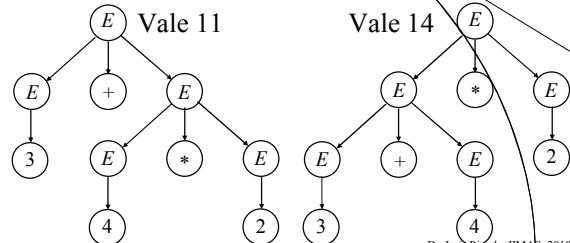
- La expresión final es la misma:
 - $E \Rightarrow^* E + E * E$
 - $E \Rightarrow^* E + E * E$
- Sus estructuras sintácticas son diferentes.



Dr. Luis Pineda, IIMAS, 20107

Una gramática ambigua

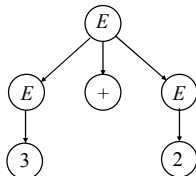
- La diferencia es significativa:
 - $E \Rightarrow E + E \Rightarrow E + E * E \Rightarrow^* 3 + 4 * 2$
 - $E \Rightarrow E * E \Rightarrow E + E * E \Rightarrow^* 3 + 4 * 2$



Dr. Luis Pineda, IIMAS, 20108

¿De quién es el "defecto"?

- Derivaciones diferentes pueden tener la misma estructura:
 - $E \Rightarrow E + E \Rightarrow 3 + E \Rightarrow 3 + 2$
 - $E \Rightarrow E + E \Rightarrow E + 2 \Rightarrow 3 + 2$



- La ambigüedad surge cuando hay más de una estructura sintáctica para la misma expresión!

Dr. Luis Pineda, IIMAS, 20109

Gramáticas ambiguas

- Una CFG $G = (V, \Sigma, S, P)$ es ambigua si existe cuando menos una cadena w en Σ^* para la cual hay más de una árbol de parseo o estructura sintáctica, cada una de éstas con raíz S y cosecha w
- Si toda cadena en el lenguaje de la gramática tiene cundo más un árbol de parseo, la gramática no es ambigua

Dr. Luis Pineda, IIMAS, 201010

Gramáticas ambiguas

- Si G es una *GLC* ambigua, tal que $L = L(G)$ & existe una G_i no ambigua tal que $L = L(G_i)$, podemos eliminar la ambigüedad reemplazando a G por G_i (usando a G_i en vez de G)

Dr. Luis Pineda, IIMAS, 201011

Eliminando la ambigüedad

- En general, no existe un algoritmo para eliminar la ambigüedad
- Hay *LLC* que sólo tienen gramáticas ambiguas!

Dr. Luis Pineda, IIMAS, 201012

Eliminando la ambigüedad

- En la práctica y para algunas aplicaciones (e.g. definir *GLC* para lenguajes de programación), es posible eliminar la ambigüedad
- Para esto es necesario estudiar las causas de la ambigüedad (específicas para una gramática ambigua dada) y proporcionar una gramática alternativa no ambigua!

Dr. Luis Pineda, IIMAS 201013

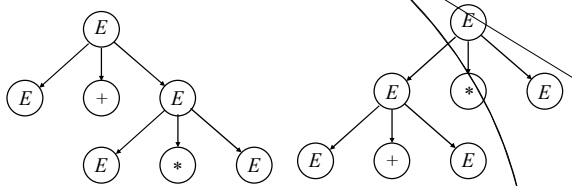
Causas de la ambigüedad

- Consideremos G_{exp} :
- Fuente 1:
 - La precedencia de los operadores no se respeta
- Fuente 2:
 - Una secuencia del mismo operador se puede agrupar tanto por la izquierda como por la derecha
 - ¡Esto no afecta si el operador es asociativo!

Dr. Luis Pineda, IIMAS 201014

Una gramática ambigua

- La precedencia de operadores no se respeta:
 - $E \Rightarrow^* E + E * E$
 - $E \Rightarrow^* E + E * E$

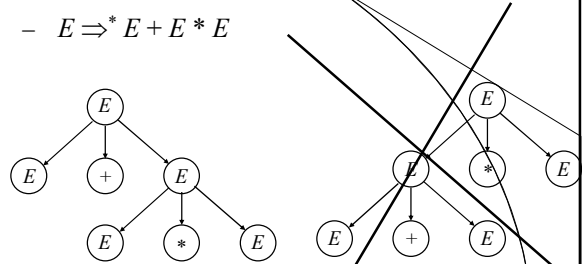


“*” tiene mayor precedencia “+” tiene mayor precedencia

Dr. Luis Pineda, IIMAS 201015

Una gramática ambigua

- Forzar la precedencia: Nos quedamos con sólo un árbol:
 - $E \Rightarrow^* E + E * E$

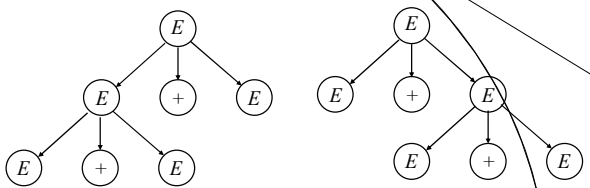


“*” tiene mayor precedencia “+” tiene mayor precedencia

Dr. Luis Pineda, IIMAS 201016

Una gramática ambigua

- Agrupaciones arbitrarias de operadores con la misma precedencia:
 - $E \Rightarrow^* E + E + E$



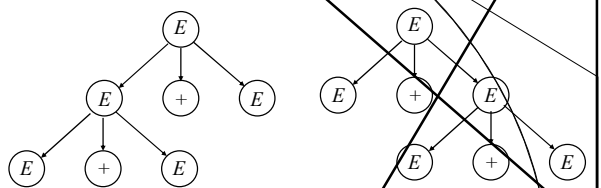
Izquierda: $E + E + E$

Derecha: $E + E + E$

Dr. Luis Pineda, IIMAS 201017

Una gramática ambigua

- Adoptamos una convención de agrupación (e.g. el mismo operador tiene mayor precedencia por la izquierda)
 - $E \Rightarrow^* E + E + E$



Izquierda: $E + E + E$

Derecha: $E + E + E$

Dr. Luis Pineda, IIMAS 201018

Eliminando la ambigüedad

- Para forzar la precedencia:
 - Usar diferentes tipos de variables para representar expresiones con el mismo *grado de atracción*
- *Binding strength*:
 - Factores
 - Términos
 - Expresiones

Dr. Luis Pineda, IIMAS 201019

Eliminando la ambigüedad

- Considerar G_{exp} (adicionando id's):
 - $G_{exp} = (\{E, I\}, \{+, *, (,), a, b, 0, 1\}, E, P)$
donde $P = \{E \rightarrow E + E \mid E * E \mid (E) \mid I,$
 $I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1\}$
- ¿Cuáles son los factores, los términos y las expresiones en G_{exp} ?

Dr. Luis Pineda, IIMAS 201020

Factores

- Una expresión que no puede partirse por ningún operador adjacente:
 - Los identificadores:
 $a1+b00*aa10 \Rightarrow (a1+b0)(0*aa10)?$
 - Expresiones entre paréntesis: Los paréntesis hacen a las expresiones que encierran unidad coherentes indestructibles! Son el mecanismo sintáctico para crear factores!

Dr. Luis Pineda, IIMAS 201021

Términos

- Una expresión que no puede partirse por el operador +:
 - ✓ $a1*a*b \Rightarrow (a1*a)*b$
 - × $a1+a*b \Rightarrow (a1+a)*b$
- $a*b$ es un término!

Dr. Luis Pineda, IIMAS 201022

El grado de atracción

- Una cadena bien formada que puede partirse por un + o un * adjacente
- La suma de dos términos

Dr. Luis Pineda, IIMAS 201023

Forzando la precedencia

- Con las variables de precedencia podemos encontrar una gramática alternativa no ambigua para G_{exp} :
 $G_{exp-1} = (\{E, T, F, I\}, \{+, *, (,), a, b, 0, 1\}, E, P)$
donde $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1,$
 $F \rightarrow I \mid (E)$
 $T \rightarrow F \mid T * F$
 $E \rightarrow T \mid E + T\}$

Dr. Luis Pineda, IIMAS 201024

Eliminando la ambigüedad

- Las producciones se diseñan de forma tal que las variables con menor grado de atracción dominen a las variables con mayor grado de atracción (i.e. las expresiones aparecen más arriba que los términos, y estos más arriba que los factores, en el árbol de parseo).

Dr. Luis Pineda, IIMAS 201025

Generar: $a + a * a$

- Las producciones:

$$P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid \Pi, \quad (E)$$

$$F \rightarrow I \mid (E)$$

$$T \rightarrow F \mid (T * F)$$

$$E \rightarrow T \mid E + T \}$$

- La derivación:

$$E \Rightarrow$$



Dr. Luis Pineda, IIMAS 201026

Generar: $a + a * a$

- Las producciones:

$$P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid \Pi,$$

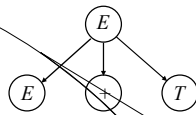
$$F \rightarrow I \mid (E)$$

$$T \rightarrow F \mid (T * F)$$

$$E \rightarrow T \mid E + T \}$$

- La derivación:

$$E \Rightarrow E + T$$



Dr. Luis Pineda, IIMAS 201027

Generar: $a + a * a$

- Las producciones:

$$P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid \Pi,$$

$$F \rightarrow I \mid (E)$$

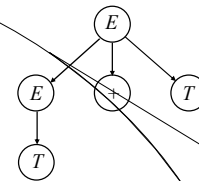
$$T \rightarrow F \mid (T * F)$$

$$E \rightarrow T \mid E + T \}$$

- La derivación:

$$E \Rightarrow E + T$$

$$\Rightarrow T + T$$



Dr. Luis Pineda, IIMAS 201028

Generar: $a + a * a$

- Las producciones:

$$P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid \Pi,$$

$$F \rightarrow I \mid (E)$$

$$T \rightarrow F \mid (T * F)$$

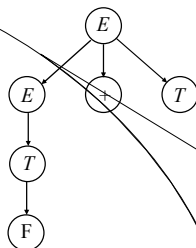
$$E \rightarrow T \mid E + T \}$$

- La derivación:

$$E \Rightarrow E + T$$

$$\Rightarrow T + T$$

$$\Rightarrow F + T$$



Dr. Luis Pineda, IIMAS 201029

Generar: $a + a * a$

- Las producciones:

$$P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid \Pi,$$

$$F \rightarrow I \mid (E)$$

$$T \rightarrow F \mid (T * F)$$

$$E \rightarrow T \mid E + T \}$$

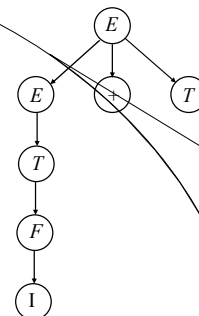
- La derivación:

$$E \Rightarrow E + T$$

$$\Rightarrow T + T$$

$$\Rightarrow F + T$$

$$\Rightarrow I + T$$



Dr. Luis Pineda, IIMAS 201030

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow E + T$
 $\Rightarrow T + T$
 $\Rightarrow F + T$
 $\Rightarrow I + T$
 $\Rightarrow a + T$

Dr. Luis Pineda, IIMAS 201031

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow E + T$
 $\Rightarrow T + T$
 $\Rightarrow F + T$
 $\Rightarrow I + T$
 $\Rightarrow a + T$
 $\Rightarrow a + (T * F)$

Dr. Luis Pineda, IIMAS 201032

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow * a + (F * F)$

Dr. Luis Pineda, IIMAS 201033

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$

Dr. Luis Pineda, IIMAS 201034

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$
 $\Rightarrow a + (a * F)$

Dr. Luis Pineda, IIMAS 201035

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid I0 \mid I1, F \rightarrow I \mid (E), T \rightarrow F \mid (T * F), E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$
 $\Rightarrow a + (a * F)$
 $\Rightarrow a + (a * I)$

Dr. Luis Pineda, IIMAS 201036

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid IO\},$
 $F \rightarrow I \mid (E)$
 $T \rightarrow F \mid (T * F)$
 $E \rightarrow T \mid E + T\}$
- La derivación:
 $E \Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$
 $\Rightarrow a + (a * F)$
 $\Rightarrow a + (a * I)$
 $\Rightarrow a + (a * a)$

Dr. Luis Pineda, IIMAS 201037

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid IO\},$
 $F \rightarrow I \mid (E)$
 $T \rightarrow F \mid (T * F)$
 $E \rightarrow T \mid E + T\}$
- La gramática no es ambigua: Las variables de menor precedencia se introducen antes, y las variables de mayor precedencia no pueden ser partidas por variables de menor precedencias, que están arriba en el árbol!

Dr. Luis Pineda, IIMAS 201038

Generar: $a + a * a$

- Las producciones:
 $P = \{I \rightarrow a \mid b \mid Ia \mid Ib \mid IO \mid IO\},$
 $F \rightarrow I \mid (E)$
 $T \rightarrow F \mid (T * F)$
 $E \rightarrow T \mid E + T\}$
- En particular, los términos y las expresiones se ramifican siempre por la izquierda!

Dr. Luis Pineda, IIMAS 201039

Gramáticas no ambiguas

Si una gramática no es ambigua sus derivaciones más izquierdas y más derechas son únicas!

- Derivación más izq:
 $E \Rightarrow E + T$
 $\Rightarrow T + T$
 $\Rightarrow F + T$
 $\Rightarrow I + T$
 $\Rightarrow a + T$
 $\Rightarrow a + (T * F)$
 $\Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$
 $\Rightarrow a + (a * F)$
 $\Rightarrow a + (a * I)$
 $\Rightarrow a + (a * a)$
- Derivación más der:
 $E \Rightarrow E + T$
 $\Rightarrow E + (T * F)$
 $\Rightarrow E + (I * a)$
 $\Rightarrow E + (T * a)$
 $\Rightarrow E + (F * a)$
 $\Rightarrow E + (I * a)$
 $\Rightarrow E + (a * a)$
 $\Rightarrow T + (a * a)$
 $\Rightarrow F + (a * a)$
 $\Rightarrow I + (a * a)$
 $\Rightarrow a + (a * a)$

Dr. Luis Pineda, IIMAS 201040

Ambigüedad y derivaciones más izquierdas

- Derivación más izq:
 $E \Rightarrow E + T$
 $\Rightarrow T + T$
 $\Rightarrow F + T$
 $\Rightarrow I + T$
 $\Rightarrow a + T$
 $\Rightarrow a + (T * F)$
 $\Rightarrow a + (F * F)$
 $\Rightarrow a + (I * F)$
 $\Rightarrow a + (a * F)$
 $\Rightarrow a + (a * I)$
 $\Rightarrow a + (a * a)$
- Teorema: Para toda gramática $G = (V, T, S, P)$ & cadenas w en T^* , w tiene dos árboles de parseo distintos si y sólo si tiene dos derivaciones más izquierdas a partir de S
- Prueba: si no fuera el caso una variable más izquierda se podría expandir en más de una forma!

Dr. Luis Pineda, IIMAS 201041

Ambigüedad Inherente

- Un lenguaje L es inherentemente ambiguo si todas sus gramáticas son ambiguas; si existe cuando menos una gramática no ambigua para L , L no es ambiguo.
 - El lenguaje de las expresiones no es ambiguo
 - Las expresiones regulares no son ambiguas

Dr. Luis Pineda, IIMAS 201042

Ambigüedad Inherente

- Ejemplo de un lenguaje inherentemente ambiguo:

$$L = \{a^n b^n c^m d^m \mid n \geq 1, m \geq 1\} \cup \{a^n b^m c^m d^n \mid n \geq 1, m \geq 1\}$$

- L es un LLC:

$$\begin{aligned} S &\rightarrow AB \mid C \\ A &\rightarrow aAb \mid ab & C &\rightarrow aCd \mid aDd \\ B &\rightarrow cBd \mid cd & D &\rightarrow bDc \mid bc \end{aligned}$$

Dr. Luis Pineda, IIMAS 201043

Ambigüedad Inherente

- L es un LLC:

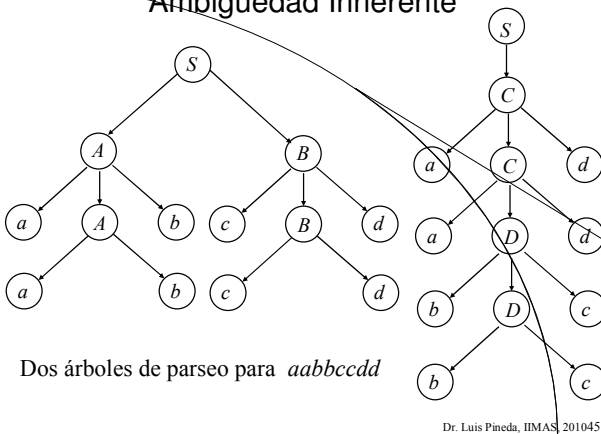
$$\begin{aligned} S &\rightarrow AB \mid C \\ A &\rightarrow aAb \mid ab & C &\rightarrow aCd \mid aDd \\ B &\rightarrow cBd \mid cd & D &\rightarrow bDc \mid bc \end{aligned}$$

- La gramática es ambigua: hay cadenas con más de una derivación más izquierda:

- Considere: $aabbccdd$ ($m = n = 2$)
- $S \Rightarrow AB \Rightarrow aAbB \Rightarrow aabbB \Rightarrow aabbcBd \Rightarrow aabbccdd$
- $S \Rightarrow C \Rightarrow aCd \Rightarrow aaDdd \Rightarrow aabDcdd \Rightarrow aabbccdd$

Dr. Luis Pineda, IIMAS 201044

Ambigüedad Inherente



Dr. Luis Pineda, IIMAS 201045

Ambigüedad Inherente

- El lenguaje:

$$L = \{a^n b^n c^m d^m \mid n \geq 1, m \geq 1\} \cup \{a^n b^m c^m d^n \mid n \geq 1, m \geq 1\}$$

- La gramática

$$\begin{aligned} S &\rightarrow AB \mid C \\ A &\rightarrow aAb \mid ab & C &\rightarrow aCd \mid aDd \\ B &\rightarrow cBd \mid cd & D &\rightarrow bDc \mid bc \end{aligned}$$

- ¿Por qué todas las gramáticas para este lenguaje son ambiguas?

- Considere cualquier cadena con $m = n$
- ¡Siempre habrá dos derivaciones para estas cadenas!

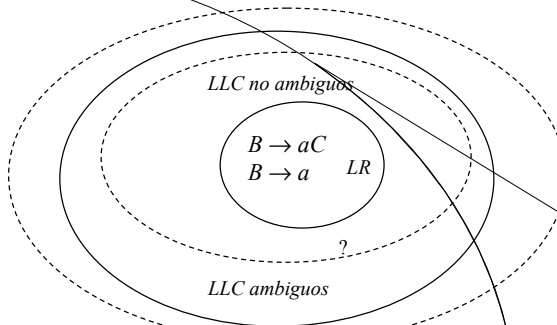
Dr. Luis Pineda, IIMAS 201046

Ambigüedad Inherente

- ¿Qué cambios podríamos hacer en la gramática?
- El problema: La disjunción (de nuevo!)
 - No hay otro mecanismo para contar el mismo número de a 's & b 's, y al mismo tiempo, para contar el mismo número de a 's & d 's
 - Lo mismo para contar las c 's & d 's &, al mismo tiempo, las b 's & c 's

Dr. Luis Pineda, IIMAS 201047

La clase de los LLC no-ambiguos



- No existe algoritmo para determinar si un lenguaje es ambiguo
- No hay manera de decidir si un lenguaje es inherentemente ambiguo!

Dr. Luis Pineda, IIMAS 201048