



# Compiladores

## Gramáticas de atributos

Artur Pereira <artur@ua.pt>,  
Miguel Oliveira e Silva <mos@ua.pt>

DETI, Universidade de Aveiro

Ano letivo de 2022-2023

## Sumário

- ① Conteúdo semântico
- ② Gramática de atributos
- ③ Avaliação dirigida pela sintaxe

## Conteúdo semântico

Ilustração com uma expressão aritmética

- Considere a gramática seguinte, onde `num` é um *token* que representa

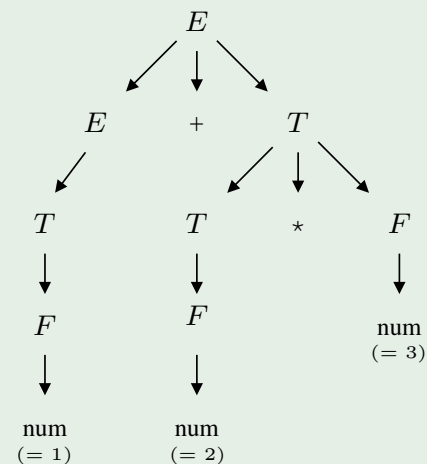
$$E \rightarrow E + T \mid T$$

um número  $T \rightarrow T * F \mid F$

$$F \rightarrow \text{num} \mid ( E )$$

- Desenhe-se a árvore de derivação da palavra "`1+2*3`"
- Como dar significado a esta árvore?

- 1 Associando a cada símbolo um atributo que armazene o valor que a sub-árvore de que é raiz representa
- 2 Relacionando os atributos associados aos símbolos de cada produção através de regras de cálculo



## Conteúdo semântico

Ilustração com uma expressão aritmética

- Considere a gramática seguinte, onde `num` é um *token* que representa

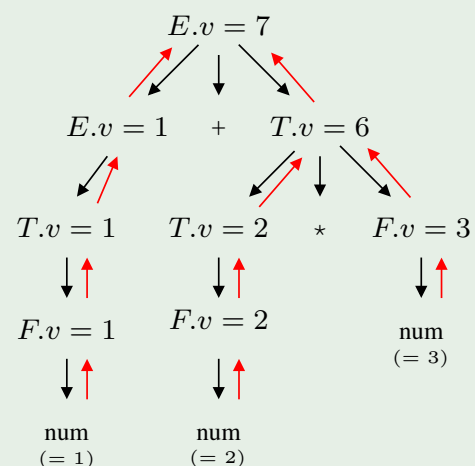
$$E \rightarrow E + T \mid T$$

um número  $T \rightarrow T * F \mid F$

$$F \rightarrow \text{num} \mid ( E )$$

- Desenhe-se a árvore de derivação da palavra "`1+2*3`"
- Como dar significado a esta árvore?

- 1 Associando a cada símbolo um atributo que armazene o valor que a sub-árvore de que é raiz representa
- 2 Relacionando os atributos associados aos símbolos de cada produção através de regras de cálculo



- As setas vermelhas representam dependência entre atributos
  - o sentido indica qual influencia qual

## Conteúdo semântico

### Ilustração com uma declaração de variáveis

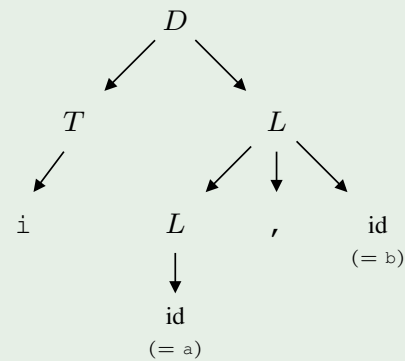
- Considere a gramática seguinte, onde *id* é um *token* que representa o nome de uma variável

$$D \rightarrow T L$$

$$T \rightarrow i \mid f$$

$$L \rightarrow id \mid L , id$$

- desenhe-se a árvore de derivação da palavra *i a, b*
- Associe-se
  - a *T* e *L* um atributo *t* que armazene o tipo



## Conteúdo semântico

### Ilustração com uma declaração de variáveis

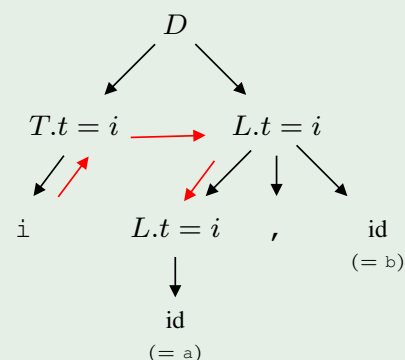
- Considere a gramática seguinte, onde *id* é um *token* que representa o nome de uma variável

$$D \rightarrow T L$$

$$T \rightarrow i \mid f$$

$$L \rightarrow id \mid L , id$$

- desenhe-se a árvore de derivação da palavra *i a, b*
- Associe-se
  - a *T* e *L* um atributo *t* que armazene o tipo



- As setas vermelhas representam dependência entre atributos
  - o sentido indica qual influencia qual

# Gramática de atributos

## Atributos, regras semânticas e definição semântica

- A análise sintática *per se* não atribui um significado às produções de uma gramática
  - É esse o papel da *gramática de atributos*
  - Isso é feito através de **atributos** e de **regras semânticas**
- Os atributos estão associados aos símbolos da gramática (terminais ou não terminais)
  - Cada símbolo terminal ou não terminal pode ter associado um conjunto de zero ou mais atributos
  - Um atributo pode ser uma palavra, um número, um tipo, uma posição de memória, ...
- As regras semânticas estão associadas às produções da gramática
  - Determinam os valores de atributos de símbolos não terminais em função de outros atributos
  - Podem ter efeitos laterais (alteração de uma estrutura de dados, ...)
- Uma **definição semântica** é composta por
  - uma gramática independente de contexto
  - um conjunto de atributos associados aos seus símbolos
  - um conjunto de regras semânticas associadas às suas produções
- Usar-se-á com o mesmo significado o termo **gramática de atributos**

# Gramática de atributos

## Regras semânticas

Seja  $G = (T, N, S, P)$  uma gramática independente do contexto

- A cada produção  $A \rightarrow B_1 B_2 \cdots B_n \in P$ , com  $B_i \in (T \cup N)^*$ , podem associar-se regras semânticas para o cálculo dos valores dos atributos de símbolos não terminais

$$b = f(c_1, c_2, \cdots, c_n)$$

onde

- $b$  é um atributo do símbolo  $A$  ou de um dos símbolos não terminais presentes em  $B_1 B_2 \cdots B_n$
  - $c_1, c_2, \cdots, c_n$  são atributos dos símbolos que ocorrem na produção
- Podem ainda associar-se regras semânticas com efeitos colaterais

$$g(c_1, c_2, \cdots, c_n)$$

- Embora este caso possa considerar-se o anterior atuando sobre um atributo fictício

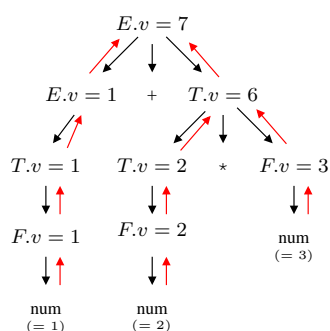
# Gramática de atributos

## Tipos de atributos

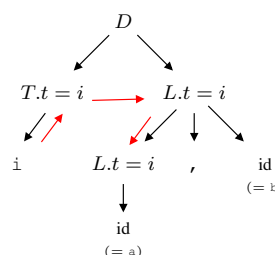
- Os atributos podem ser classificados como **sintetizados** ou **herdados**
- Considere-se uma produção  $A \rightarrow B_1 B_2 \cdots B_n \in P$ , com  $B_i \in (T \cup N)^*$ , e uma função de cálculo de um atributo associada a essa produção

$$b = f(c_1, c_2, \dots, c_n)$$

- O atributo  $b$  diz-se **sintetizado** se  $b$  está associado a  $A$  e todos os  $c_j$ , com  $j = 1, 2, \dots, n$ , estão associados a símbolos do corpo da produção
- O atributo  $b$  diz-se **herdado** se  $b$  está associado a um dos símbolos não terminais do corpo da produção



- Todos os atributos são sintetizados



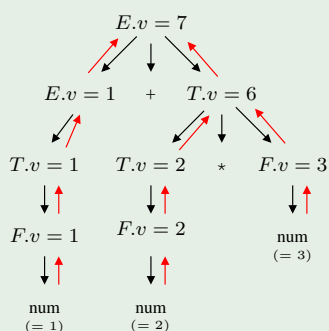
- $T.t$  é sintetizado
- $L.t$  é herdado

# Gramática de atributos

## Representação

- Uma gramática de atributos pode ser representada por uma tabela em que se associam as regras semânticas às produções da gramática

- Para o exemplo das expressões aritméticas, tem-se



Produções	Regras semânticas
$F \rightarrow \text{num}$	$F.v = \text{num}.v$
$F \rightarrow ( E )$	$F.v = E.v$
$T \rightarrow F$	$T.v = F.v$
$T_1 \rightarrow T_2 * F$	$T_1.v = T_2.v * F.v$
$E \rightarrow T$	$E.v = T.v$
$E_1 \rightarrow E_2 + T$	$E_1.v = E_2.v + T.v$

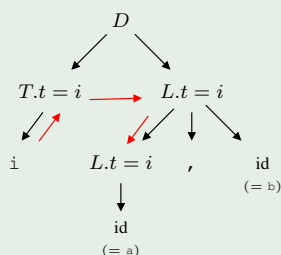
- Note que se assume que o símbolo terminal `num` tem um atributo chamado  $v$  com o valor correspondente.
- O ANTLR não suporta atributos nos terminais (*tokens*)

# Gramática de atributos

## Representação

- Uma gramática de atributos pode ser representada por uma tabela em que se associam as regras semânticas às produções da gramática

- Para o exemplo da declaração de variáveis, tem-se



Produções	Regras semânticas
$T \rightarrow i$	$T.t = \text{int}$
$T \rightarrow f$	$T.t = \text{float}$
$D \rightarrow T L$	$L.t = T.t$
$L_1 \rightarrow L_2 , \text{id}$	$L_2.t = L_1.t$ $\text{addsym}(\text{id}.n, L_1.t)$
$L \rightarrow \text{id}$	$\text{addsym}(\text{id}.n, L.t)$

- Assume-se que o símbolo terminal `id` tem um atributo chamado `n` com o valor correspondente
- Neste caso, para além do cálculo de atributos, faz-se a inserção numa tabela de símbolos (`addsym`)

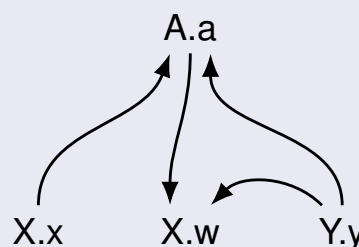
## Avaliação dirigida pela sintaxe

- Numa **avaliação dirigida pela sintaxe** o cálculo dos atributos é feito à medida que é feita a análise sintática.
- Num analisador sintático ascendente (caso do bison) todos os atributos têm de ser sintetizados
- Num analisador sintático descendente (caso do Antlr) além de sintetizados os atributos podem ser herdados, desde que de símbolos à esquerda ou do símbolo pai
- para definir a ordem de cálculo dos atributos, usa-se o **grafo de dependências**

$$A \rightarrow X Y$$

$$A.a = f(X.x, Y.y)$$

$$X.w = g(A.a, Y.y)$$



- Aqui as setas apontam no sentido das dependências