

# Arith ++

Sintaxe, semântica e tipagem

---

João Reis  
joao.reis@ubi.pt

## 1 Introdução

Este documento serve de referência para o resultado esperado na primeira entrega do trabalho prático desta unidade curricular. Aqui é descrita uma proposta de extensão à linguagem `Arith` chamada `Arith++`, especificando em toda a sua completude a sintaxe abstrata, semântica operacional *small-steps* e tipagem.

## 2 Sintaxe Abstrata

### 2.1 Constantes

A linguagem `Arith++` aceita constantes numéricas inteiras e de vírgula flutuante, assim como constantes *booleanas* e vetores vazios para os tipos mencionados anteriormente.

```
 $\langle c \rangle ::=$  inteiro (0, 5, 123, ...)  
| float (0.01, .3, 2.1, 100., ...)  
| booleano (true, false)  
|  $\langle t \rangle$  [ ] (vetor vazio de um determinado tipo)
```

```
 $\langle t \rangle ::=$  int  
| float  
| bool
```

### 2.2 Expressões

Uma expressão em `Arith++` assume as seguintes formas:

- uma constante;
- uma variável

- o acesso a uma posição de um vetor;
- uma expressão dentro de parênteses;
- operações unárias aritméticas e *booleanas*;
- operações binárias aritméticas, *booleanas* e de comparação;
- atribuição local de variáveis;
- chamada de funções.

```

⟨e⟩ ::= ⟨c⟩
      | ⟨ident⟩
      | ⟨ident⟩ . ( ⟨e⟩ )
      | ( ⟨e⟩ )
      | ⟨e⟩ + ⟨e⟩
      | ⟨e⟩ - ⟨e⟩
      | ⟨e⟩ * ⟨e⟩
      | ⟨e⟩ / ⟨e⟩
      | - ⟨e⟩
      | ⟨e⟩ && ⟨e⟩
      | ⟨e⟩ || ⟨e⟩
      | ! ⟨e⟩
      | ⟨e⟩ > ⟨e⟩
      | ⟨e⟩ < ⟨e⟩
      | ⟨e⟩ >= ⟨e⟩
      | ⟨e⟩ <= ⟨e⟩
      | ⟨e⟩ == ⟨e⟩
      | let ⟨ident⟩ = ⟨e⟩ in ⟨e⟩
      | ⟨ident⟩ ⟨e⟩

```

## 2.3 Instruções

Um programa em `Arith++` é composto por uma ou mais instruções (*statements*). O conjunto de instruções desta linguagem é composto por:

- a atribuição de variáveis globais;
- impressão de uma expressão no ecrã;
- estruturas condicionais e de ciclo *if then else*, *while* e *for* (este último com duas variantes, percorrer os elementos de um vetor ou percorrer um intervalo de números inteiros);
- atribuição a uma posição de um vetor;
- declaração de funções;
- não fazer nada;
- sequência de instruções.

$\langle s \rangle ::= \text{set } \langle \text{ident} \rangle = \langle e \rangle$   $\text{print } \langle e \rangle$   $\text{if } \langle e \rangle \text{ then } \langle s \rangle [\text{else } \langle s \rangle] \text{ done}$   $\text{while } \langle e \rangle \text{ do } \langle s \rangle \text{ done}$   $\text{for } \langle \text{ident} \rangle \text{ in } \langle e \rangle \text{ do } \langle s \rangle \text{ done}$   $\text{for } \langle \text{ident} \rangle = \langle e \rangle \text{ to } \langle e \rangle \text{ do } \langle s \rangle \text{ done}$   $\langle e \rangle . ( \langle e \rangle ) <- \langle e \rangle$	$\text{func } \langle \text{ident} \rangle ( \langle \text{params} \rangle ) \text{ begin } \langle e \rangle \text{ end}$   $\text{skip}$   $\langle s \rangle \langle s \rangle$  $\langle \text{params} \rangle ::= \langle t \rangle \langle \text{ident} \rangle$   $\langle t \rangle \langle \text{ident} \rangle , \langle \text{params} \rangle$
---	--

### 3 Semântica

De seguida, definimos a semântica operacional *small-steps* para a linguagem *Arith++*. Definimos como valores desta linguagem os seguintes:

$v ::= i$  (valor inteiro)  
 $f$  (valor vírgula flutuante)  
 $b$  (valor booleano)  
 $x$  (variável vista como um endereço)  
 $[v_0, \dots, v_n]$  (vetor de valores)  
 $\text{fun } (t_1 x_1, \dots, t_n x_n) \Rightarrow e$  (função)

#### 3.1 Expressões

A avaliação das expressões desta linguagem termina sempre. Desta forma, optamos por definir a semântica das expressões com recurso à semântica *big-steps*. Os casos não-triviais serão acompanhados de uma breve descrição.

Aqui temos a semântica das constantes:

$$\overline{E, i \rightarrow i} \quad \overline{E, f \rightarrow f} \quad \overline{E, b \rightarrow b} \quad \overline{E, t [] \rightarrow []} \quad \overline{E, id \rightarrow E(id)}$$

De seguida, definimos a semântica do acesso a uma posição de um vetor:

$$\frac{\overline{E, id \rightarrow x} \quad \overline{E, e \rightarrow v}}{\overline{E, id.(e) \rightarrow E(x[v])}}$$

A variável *id* avalia para um endereço de memória *x* e a expressão *e* avalia para um valor *v*. A expressão *x.(e)* avalia desta forma para o valor localizado na memória no endereço  $x + v$  ( $x[v]$ ).

Definimos de forma trivial as operações binárias e unárias desta linguagem:

$$\frac{\overline{E, e \rightarrow i}}{\overline{E, -e \rightarrow -i}} \quad \frac{\overline{E, e_1 \rightarrow i_1} \quad \overline{E, e_2 \rightarrow i_2} \quad i = i_1 + i_2}{\overline{E, e_1 + e_2 \rightarrow i}} \quad \frac{\overline{E, e_1 \rightarrow i_1} \quad \overline{E, e_2 \rightarrow i_2} \quad i = i_1 - i_2}{\overline{E, e_1 - e_2 \rightarrow i}}$$

$$\frac{\overline{E, e_1 \rightarrow i_1} \quad \overline{E, e_2 \rightarrow i_2} \quad i = i_1 \times i_2}{\overline{E, e_1 * e_2 \rightarrow i}} \quad \frac{\overline{E, e_1 \rightarrow i_1} \quad \overline{E, e_2 \rightarrow i_2} \quad i_2 \neq 0 \quad i = i_1 \div i_2}{\overline{E, e_1 / e_2 \rightarrow i}}$$

$$\frac{\overline{E, e_1 \rightarrow false}}{\overline{E, e_1 \ \&\& \ e_2 \rightarrow false}} \quad \frac{\overline{E, e_1 \rightarrow true} \quad \overline{E, e_2 \rightarrow false}}{\overline{E, e_1 \ \&\& \ e_2 \rightarrow false}} \quad \frac{\overline{E, e_1 \rightarrow true} \quad \overline{E, e_2 \rightarrow true}}{\overline{E, e_1 \ \&\& \ e_2 \rightarrow true}}$$

$$\begin{array}{c}
\frac{E, e_1 \rightarrow true}{E, e_1 \ || \ e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow false \quad E, e_2 \rightarrow true}{E, e_1 \ || \ e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow false \quad E, e_2 \rightarrow false}{E, e_1 \ || \ e_2 \rightarrow false} \\
\\
\frac{E, e \rightarrow false}{E, !e \rightarrow true} \quad \frac{E, e \rightarrow true}{E, !e \rightarrow false} \\
\\
\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 > v_2}{E, e_1 > e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 \leq v_2}{E, e_1 > e_2 \rightarrow false} \\
\\
\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 \geq v_2}{E, e_1 \geq e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 < v_2}{E, e_1 \geq e_2 \rightarrow false} \\
\\
\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 < v_2}{E, e_1 < e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 \geq v_2}{E, e_1 < e_2 \rightarrow false} \\
\\
\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 \leq v_2}{E, e_1 \leq e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 > v_2}{E, e_1 \leq e_2 \rightarrow false} \\
\\
\frac{E, e_1 \rightarrow v \quad E, e_2 \rightarrow v}{E, e_1 == e_2 \rightarrow true} \quad \frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 \neq v_2}{E, e_1 == e_2 \rightarrow false}
\end{array}$$

A semântica da expressão **let in** é definida da seguinte forma:

$$\frac{E, e_1 \rightarrow v_1 \quad E\{x \mapsto v_1\}, e_2 \rightarrow v_2}{E, \text{let } x = e_1 \text{ in } e_2 \rightarrow v_2}$$

A expressão  $e_1$  é avaliada no ambiente  $E$  para o valor  $v_1$ . De seguida, avalia-se a expressão  $e_2$  no ambiente  $E$ , acrescido da variável  $x$  que tem o valor  $v_1$ , para  $v_2$ .

$$\frac{E, e \rightarrow fun(x_1, \dots, x_n) \Rightarrow e_f \quad E, e_1 \rightarrow v_1 \ \dots \ E, e_n \rightarrow v_n \quad E\{x_1 \mapsto v_1, \dots, x_n \mapsto v_n\}, e_f \rightarrow v}{E, e(e_1, \dots, e_n) \rightarrow v}$$

### 3.2 Instruções

Definimos a seguir a semântica das instruções da linguagem **Arith++** usando a semântica *small-steps*, mas mantendo sempre a semântica *big-steps* sempre que nos referimos às expressões.

Para a sequência de instruções, definimos dois casos:

$$\frac{}{E, \text{skip } s \rightarrow E, s} \quad \frac{E, s_1 \rightarrow E_1, s'_1}{E, s_1 \ s_2 \rightarrow E_1, s'_1 \ s_2}$$

A atribuição de variáveis globais resulta num novo ambiente em que a variável em causa irá ter associado o valor resultante da avaliação da expressão de atribuição:

$$\frac{E, e \rightarrow v}{E, \text{set } x = e \rightarrow E\{x \mapsto v\}, \text{skip}} \quad \frac{E, e \rightarrow v}{E, \text{print } e \rightarrow E, \text{skip}}$$

A instrução condicional **if** resulta no passo em que se avalia a instrução  $s$  caso a condição  $e$  se avalie para **true**, ou na avaliação da instrução **skip** caso a condição se avalie para **false**.

$$\frac{E, e \rightarrow false}{E, \text{ if } e \text{ then } s \text{ done} \rightarrow E, \text{ skip}} \quad \frac{E, e \rightarrow true}{E, \text{ if } e \text{ then } s \text{ done} \rightarrow E, s}$$

De forma semelhante, definimos a instrução condicional **if else**, mas desta vez avaliamos  $s_2$  caso a condição se avalie para **false**.

$$\frac{E, e \rightarrow false}{E, \text{ if } e \text{ then } s \text{ else } s_2 \text{ done} \rightarrow E, s_2} \quad \frac{E, e \rightarrow true}{E, \text{ if } e \text{ then } s_1 \text{ else } s_2 \text{ done} \rightarrow E, s_1}$$

Do ciclo **while** resulta no passo de avaliação da sequência de instruções composta pelo seu corpo seguido do mesmo ciclo **while** sempre que a condição se avalie para **true**. Caso a condição se avalie para **false**, resulta no passo de avaliação da instrução **skip**.

$$\frac{E, e \rightarrow true}{E, \text{ while } e \text{ do } s \text{ done} \rightarrow E, s \text{ while } e \text{ do } s \text{ done}} \quad \frac{E, e \rightarrow false}{E, \text{ while } e \text{ do } s \text{ done} \rightarrow E, \text{ skip}}$$

A instrução **for** é definida em duas variações, como descrito anteriormente. No caso da iteração sobre os elementos de um vetor, a avaliação do **for** resulta no passo de avaliação no mesmo ambiente, acrescido da variável de ciclo mapeada para o valor da cabeça do vetor, da sequência composta pelo seu corpo e novo ciclo **for** sobre a cauda do vetor a iterar. Quando o vetor a iterar está vazio, resulta no passo de avaliação da instrução **skip**.

$$\frac{E, e \rightarrow [v_0, \dots, v_n]}{E, \text{ for } x \text{ in } e \text{ do } s \text{ done} \rightarrow E\{x \mapsto v_0\}, s \text{ for } x \text{ in } [v_1, \dots, v_n] \text{ do } s \text{ done}}$$

$$\frac{E, e \rightarrow []}{E, \text{ for } x \text{ in } e \text{ do } s \text{ done} \rightarrow E, \text{ skip}}$$

Temos de seguida o caso do **for** sobre um intervalo de valores. Quando a expressão que define o limite da iteração se avalia para um valor superior ao valor para o qual se avalia a expressão que define o valor inicial, a avaliação deste caso resulta no passo de avaliação da sequência de instruções composta pelo corpo do **for** seguido de novo **for** em que a expressão inicial é o valor inicial anterior somado com 1. Quando a expressão que define o limite da iteração se avalia para um valor inferior ao valor para o qual se avalia a expressão que define o valor inicial, a avaliação deste caso resulta num passo igual ao anterior com a exceção de que se subtrai 1 ao valor inicial. Quando a expressão inicial e a expressão de limite avaliam para o mesmo valor, resulta no passo de avaliação da instrução **skip**.

$$\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 < v_2}{E, \text{ for } x = e_1 \text{ to } e_2 \text{ do } s \text{ done} \rightarrow E\{x \mapsto v_1\}, s \text{ for } x = (v_1 + 1) \text{ to } e_2 \text{ do } s \text{ done}}$$

A atribuição de uma posição de um vetor e definição de uma função comportam-se de forma semelhante à atribuição de variáveis globais.

$$\frac{E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2 \quad v_1 > v_2}{E, \text{ for } x = e_1 \text{ to } e_2 \text{ do } s \text{ done} \rightarrow E\{x \mapsto v_1\}, s \text{ for } x = (v_1 - 1) \text{ to } e_2 \text{ do } s \text{ done}}$$

$$\frac{E, e_1 \rightarrow v \quad E, e_2 \rightarrow v}{E, \text{ for } x = e_1 \text{ to } e_2 \text{ do } s \text{ done} \rightarrow E, \text{ skip}}$$

$$\frac{E, id \rightarrow x \quad E, e_1 \rightarrow v_1 \quad E, e_2 \rightarrow v_2}{E, id.(e_1) < - e_2 \rightarrow E\{x[v_1] \mapsto v_2\}, \text{skip}}$$

$$\overline{E, \text{func } x(t_1 a_1, \dots, t_n a_n) \text{ begin } e \text{ end} \rightarrow E\{x \mapsto \text{fun } x(a_1, \dots, a_n) \Rightarrow e\}, \text{skip}}$$

## 4 Tipagem

$$\begin{array}{l} \tau ::= \text{int} \\ \quad | \text{bool} \\ \quad | \text{float} \\ \quad | [\tau] \\ \quad | \tau \times \tau \times \dots \times \tau \rightarrow \tau \end{array}$$

$$\overline{\Gamma \vdash i : \text{int}} \quad \overline{\Gamma \vdash f : \text{float}} \quad \overline{\Gamma \vdash b : \text{bool}} \quad \overline{\Gamma \vdash x : \Gamma(x)}$$

$$\overline{\Gamma \vdash \text{int } [] : [\text{int}]} \quad \overline{\Gamma \vdash \text{float } [] : [\text{float}]} \quad \overline{\Gamma \vdash \text{bool } [] : [\text{bool}]} \quad \frac{\Gamma \vdash id : [\tau] \quad \Gamma \vdash e : \text{int}}{\Gamma \vdash id.(e) : \tau}$$

$$\frac{\Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \text{int}}{\Gamma \vdash e_1 + e_2 : \text{int}} \quad \frac{\Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \text{int}}{\Gamma \vdash e_1 - e_2 : \text{int}} \quad \frac{\Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \text{int}}{\Gamma \vdash e_1 * e_2 : \text{int}}$$

$$\frac{\Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \text{int}}{\Gamma \vdash e_1 / e_2 : \text{int}} \quad \frac{\Gamma \vdash e : \text{int}}{\Gamma \vdash -e : \text{int}}$$

$$\frac{\Gamma \vdash e_1 : \text{float} \quad \Gamma \vdash e_2 : \text{float}}{\Gamma \vdash e_1 + e_2 : \text{float}} \quad \frac{\Gamma \vdash e_1 : \text{float} \quad \Gamma \vdash e_2 : \text{float}}{\Gamma \vdash e_1 - e_2 : \text{float}} \quad \frac{\Gamma \vdash e_1 : \text{float} \quad \Gamma \vdash e_2 : \text{float}}{\Gamma \vdash e_1 * e_2 : \text{float}}$$

$$\frac{\Gamma \vdash e_1 : \text{float} \quad \Gamma \vdash e_2 : \text{float}}{\Gamma \vdash e_1 / e_2 : \text{float}} \quad \frac{\Gamma \vdash e : \text{float}}{\Gamma \vdash -e : \text{float}}$$

$$\frac{\Gamma \vdash e_1 : \text{bool} \quad \Gamma \vdash e_2 : \text{bool}}{\Gamma \vdash e_1 \ \&\& \ e_2 : \text{bool}} \quad \frac{\Gamma \vdash e_1 : \text{bool} \quad \Gamma \vdash e_2 : \text{bool}}{\Gamma \vdash e_1 \ || \ e_2 : \text{bool}} \quad \frac{\Gamma \vdash e : \text{bool}}{\Gamma \vdash !e : \text{bool}}$$

$$\frac{\Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 > e_2 : \text{bool}} \quad \frac{\Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 < e_2 : \text{bool}} \quad \frac{\Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 \geq e_2 : \text{bool}} \quad \frac{\Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 \leq e_2 : \text{bool}}$$

$$\frac{\Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash e_1 == e_2 : \text{bool}} \quad \frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma + x : \tau_1 \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma + x : \tau_1 \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2}$$

$$\frac{\Gamma \vdash x : \tau_1 \times \dots \times \tau_n \rightarrow \tau \quad \Gamma \vdash e_1 : \tau_1 \quad \dots \quad \Gamma \vdash e_n : \tau_n}{\Gamma \vdash x(e_1, \dots, e_n) : \tau}$$

$$\begin{array}{c}
\frac{}{\Gamma \vdash \text{skip}} \quad \frac{\Gamma \vdash s_1 \quad \Gamma \vdash s_2}{\Gamma \vdash s_1 s_2} \quad \frac{\Gamma \vdash e : \tau \quad \Gamma + id : \tau}{\Gamma \vdash \text{set } id = e} \quad \frac{\Gamma \vdash e : \tau}{\Gamma \vdash \text{print } e} \\
\\
\frac{\Gamma \vdash e : \text{bool} \quad \Gamma \vdash s}{\Gamma \vdash \text{if } e \text{ then } s \text{ done}} \quad \frac{\Gamma \vdash e : \text{bool} \quad \Gamma \vdash s_1 \quad \Gamma \vdash s_2}{\Gamma \vdash \text{if } e \text{ then } s_1 \text{ else } s_2 \text{ done}} \quad \frac{\Gamma \vdash e : [\tau] \quad \Gamma + id : \tau \quad \Gamma \vdash s}{\Gamma \vdash \text{for } id \text{ in } e \text{ do } s \text{ done}} \\
\\
\frac{\Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \text{int} \quad \Gamma + id : \text{int}}{\Gamma \vdash \text{for } id = e_1 \text{ to } e_2 \text{ do } s \text{ done}} \quad \frac{\Gamma \vdash e : \text{bool} \quad \Gamma \vdash s}{\Gamma \vdash \text{while } e \text{ do } s \text{ done}} \\
\\
\frac{\Gamma \vdash id : [\tau] \quad \Gamma \vdash e_1 : \text{int} \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash id.(e_1) <- e_2} \quad \frac{\Gamma \vdash e : \tau \quad \Gamma + id : t_1 \times \dots \times t_n \rightarrow \tau}{\Gamma \vdash \text{func } id (t_1 id_1, \dots, t_n id_n) \text{ begin } e \text{ end}}
\end{array}$$