Nomes: Gustavo Pioli Resende - 12111BCC010 Natan Gonçalves de Lyra - 12111BCC006

Vinícius Dias Martins Rocha - 12111BCC039

# **Projeto Compiladores**

# 1. Especificação Da Linguagem

# 1.1 Definição da Gramática Livre de Contexto (GLC)

Uma GLC G pode ser representada por uma quádrupla:

- V = Conjunto dos Símbolos Não-Terminais
- T = Conjunto dos Símbolos Terminais
- P = Conjunto de Produções (regras de transição)
- S = INICIO

#### **Símbolos Terminais:**

- main
- ID
- begin
- end
- int
- char
- float
- if
- then
- else
- while
- do
- repeat
- until
- número
- caractere

- letra
- relop
- ws
- {
- }
- (
- )
- :=
- \_
- •
- .
- \*
- \_ 44
- **-** +
- -
- /

Símbolos Não-Terminais ANTES da Fatoração e Remoção da Recursão à Esquerda:

- <INICIO>
- <tipo>
- <bloco>
- <declaracao variaveis>
- <declaracao variavel>
- ta ids>
- <sequencia comandos>
- <comando>
- <condição>
- <expressão>
- <comando\_selecao>
- <comando ou bloco>
- <comando repeticao>
- <comando atribuicao>
- <condicao>
- <expressao>

Produções Gramaticais (regras de transição) ANTES da Fatoração e Remoção da Recursão à Esquerda:

- <INICIO> ::= main ID <bloco>
- **<bloco>** ::= begin **<**declaração variaveis> **<**sequencia comandos> end
- <declaracao variaveis> ::= <declaracao variavel> <declaracao variaveis> | ε

- <declaracao variavel> ::= <tipo> : !:= <tipo> : !:= <tipo> : </ti>
- <tipo> ::= char | int | float
- dista ids> ::= ID | ID, ids>
- <sequencia comandos> ::= <comando> <sequencia comandos> | ε
- <comando> ::= <comando\_selecao> | <comando\_repeticao> | <comando atribuicao>
- <comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco> | if ( <condicao> ) then <comando\_ou\_bloco> else <comando\_ou\_bloco>
- <comando ou bloco> ::= <comando> | <bloco>
- <comando\_repeticao> ::= while ( <condicao> ) do <comando\_ou\_bloco> | repeat <comando\_ou\_bloco> until ( <condicao> );
- <comando atribuicao> ::= ID := <expressao> ;
- <condicao> ::= <expressao> relop <expressao>
- <expressao> ::= num | caractere | ID | (<expressao>) | <expressao> operador <expressao>

## 1.2 Definição da Gramática Livre de Contexto (GLC)

LEXEMA	TOKEN	ATRIBUTO
main	main	-
begin	begin	-
end	end	-
int	int	-
char	char	-
float	float	-
if	if	-
then	then	-
else	else	-
while	while	-
do	do	-
repeat	repeat	-
until	until	-
Qualquer número	NUM	Posição na tabela de símbolos
'Qualquer caractere'	caractere	Posição na tabela de símbolos

Qualquer ID	ID	Posição na tabela de símbolos
>	relop	GT
<	relop	LT
>=	relop	GE
<=	relop	LE
==	relop	EQ
!=	relop	NE
:=	:=	-
{	{	-
}	}	-
(	(	-
)	)	-
+	+	-
-	-	-
/	/	-
*	*	-
**	**	-
,	,	-
;	;	-
:	:	-
Texto entre chaves { }	-	-
ws	-	-

# 1.3 Definição dos padrões de cada token

```
• main
                  main
  begin →
                 begin
  end
                 end
  int
                 int
  char
                 char
  float
                 float

    if

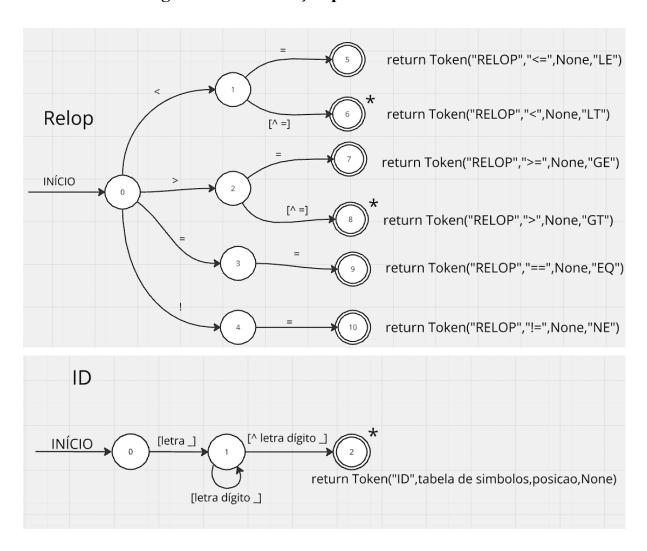
                 if
  then
                 then
• else
                 else
• while →
                 while
  do
                 do
repeat →
                 repeat
  until
                 until
                  {
                  }
                 (
                 )
                 :=
                 \backslash +
                 \*
                 \*\*
                 /
digito
                 [0-9]
• digitos →
                 digito+
                 digitos(\land digitos)?(E[\land +-]?digitos)?
• NUM →
                 [a-zA-Z]
• letra

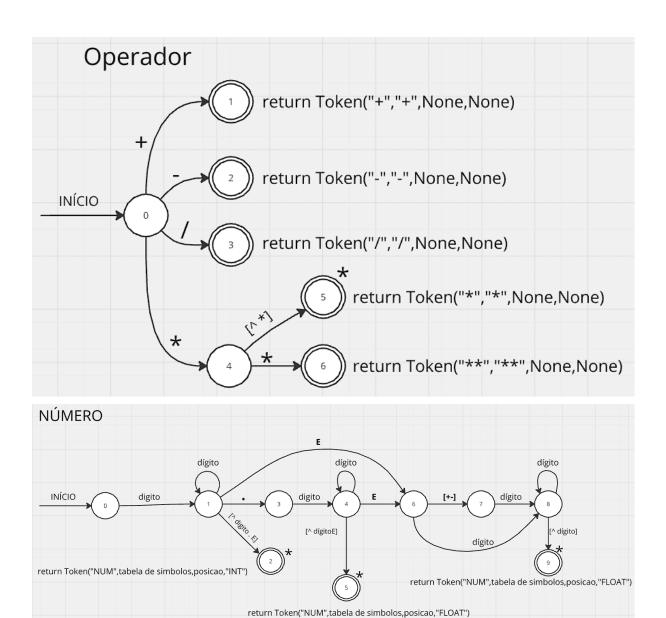
    ID

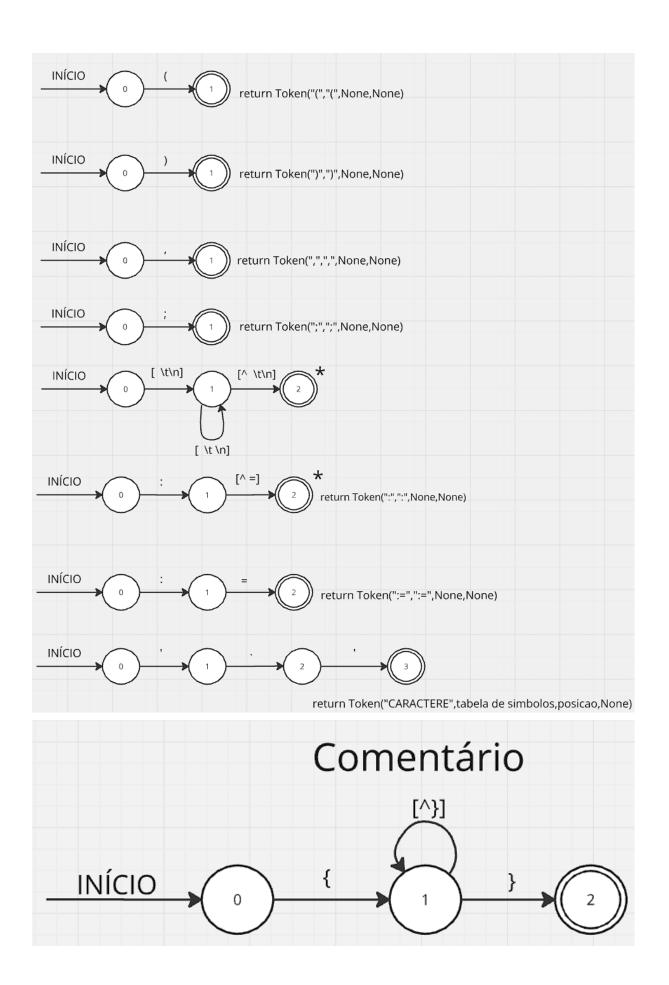
                 [letra _ ][digito letra _ ] *
• caractere →
• relop
                 < | <= | > | >= | !=
• comentário → { [^ }]* }
                 [ \t \n]
• ws
```

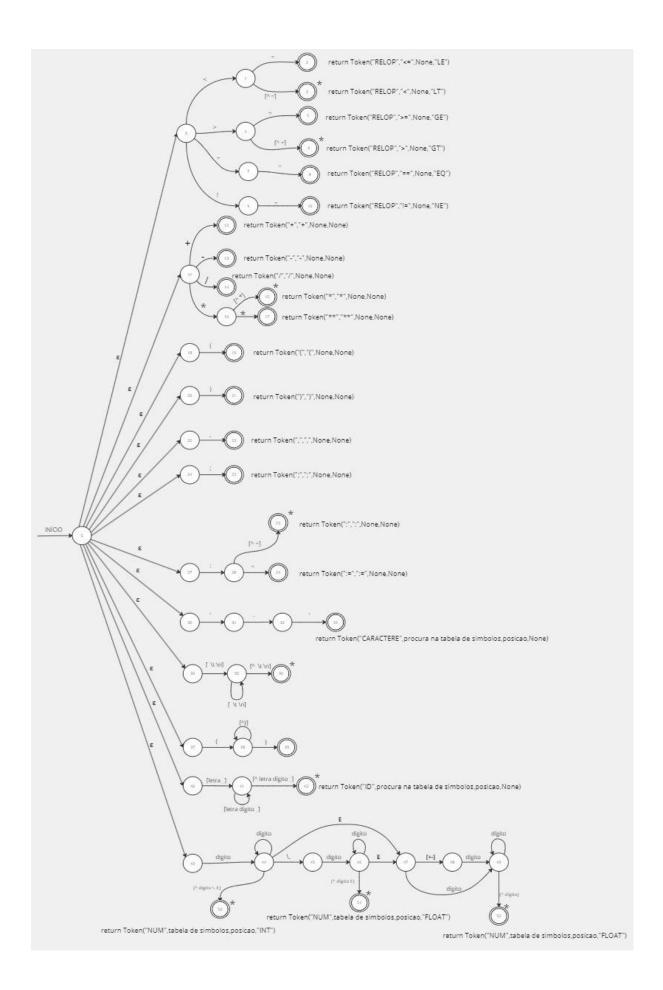
# 2. Análise Léxica

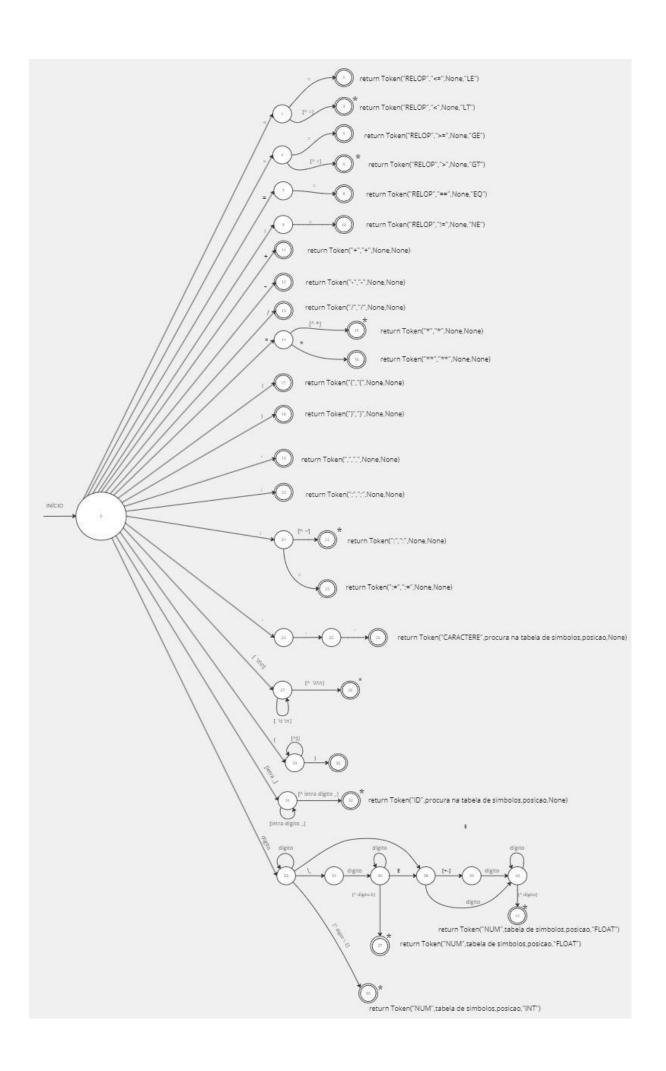
# 2.1 Diagramas de Transição para os Tokens











#### 3. Análise Sintática

#### 3.1 Ajustes da GLC para que a linguagem seja LL(1)

A gramática de produção **< lista ids>** possuía ambiguidade no **ID**:

• dista\_ids> ::= ID | ID, <lista\_ids>

Para remover a ambiguidade da gramática de produção **lista\_ids>** foram feitos os seguintes ajustes:

- dista ids>::= ID<lista ids'>
- <lista\_ids'> ::= ,ID<lista\_ids'> |  $\epsilon$

A gramática de produção **<comando selecao>** possuía ambiguidade no **if**:

• <comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco> | if ( <condicao> ) then <comando ou bloco> else <comando ou bloco>

Para remover a ambiguidade da gramática de produção **<comando\_selecao>** foram feitos os seguintes ajustes:

- <comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco> <comando\_selecao'>
- <comando\_selecao'> ::= else <comando\_ou\_bloco> |  $\epsilon$

A gramática de produção **expressao** possuía recursão à esquerda:

• <expressao> ::= num | caractere | ID | (<expressao>) | <expressao> operador <expressao>

Para remover a recursão à esquerda da gramática de produção **<expressao>** foram feitos os seguintes ajustes:

- <expressao> ::= <termo><expressao'>
- **<expressao'>** ::= + <termo> <expressao'> | <termo> <expressao'> | ε
- <termo> ::= <exp><termo'>
- <termo'>::= \* <exp> <termo'> | / <exp> <termo'> |  $\epsilon$
- **<exp>::=** <fator> <exp'>
- $\langle \exp' \rangle ::= ** \langle fator \rangle \langle \exp' \rangle \mid \varepsilon$
- <fator> ::= num | caractere | ID | (<expressao>)

# Símbolos Não-Terminais DEPOIS da Fatoração e Remoção da Recursão à Esquerda:

- <INICIO>
- <tipo>
- <bloco>

- <declaracao variaveis>
- <declaracao variavel>
- ta ids>
- ta ids'>
- <sequencia comandos>
- <comando>
- <condição>
- <expressão>
- <comando selecao>
- <comando selecao'>
- <comando ou bloco>
- <comando repeticao>
- <comando atribuicao>
- <condicao>
- <expressao>
- <expressao'>
- <termo>
- <termo'>
- <exp>
- <exp'>
- <fator>

# Produções Gramaticais (regras de transição) DEPOIS da Fatoração e Remoção da Recursão à Esquerda:

- <INICIO> ::= main ID <bloco>
- **<blood>** ::= **begin <** declaração variaveis> < sequencia comandos> **end**
- <declaração variaveis> ::= <declaração variavei> <declaração variaveis> | ε
- <declaração variavel> ::= <tipo> : !:= stipo> : <li
- <tipo> ::= char | int | float
- dista ids>::= IDids'>
- < lista ids'> ::= ,ID < lista ids'>  $\mid \varepsilon \mid$
- <sequencia comandos> ::= <comando> <sequencia comandos> | ε
- <comando> ::= <comando selecao> | <comando repeticao> | <comando atribuicao>
- <comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco> <comando\_selecao'>
- <comando selecao'> ::= else <comando ou bloco> | ε
- <comando ou bloco> ::= <comando> | <bloco>
- <comando\_repeticao> ::= while ( <condicao> ) do <comando\_ou\_bloco> | repeat <comando ou bloco> until ( <condicao> );
- <comando atribuicao> ::= ID := <expressao> ;
- <condicao> ::= <expressão> relop <expressão>
- <expressao> ::= <termo><expressao'>

- **<expressao'>** ::= + <termo> <expressao'> | <termo> <expressao'> | ε
- <termo> ::= <exp><termo'>
- <termo'>::= \* <exp> <termo'> | / <exp> <termo'> |  $\epsilon$
- <exp>::= <fator><exp'>
- $\langle \exp' \rangle ::= ** \langle fator \rangle \langle \exp' \rangle \mid \varepsilon$
- <fator> ::= num | caractere | ID | (<expressao>)

#### 3.2 Cálculo e Tabela de FIRST

- FIRST(<INICIO>) ::= FIRST(main ID <bloco>) = {main}<INICIO> ::= main ID <bloco>
- FIRST(<blood>) ::= FIRST(begin < declaracao\_variaveis> < sequencia\_comandos> end) = {begin}
- FIRST(<declaração\_variaveis>) ::= FIRST(<declaracao\_variavel>
   <declaracao\_variaveis>) U {ε} = FIRST(<declaracao\_variavel>) U {ε} = {char, int, float, ε}
  - <declaração\_variaveis> ::= <declaracao\_variavel> <declaracao\_variaveis> | ε
- FIRST(<declaração\_variavel>) ::= FIRST(<tipo> :
  - <declaração variavel> ::= <tipo> : : ids> ;
- FIRST(<tipo>) ::= FIRST(char) U FIRST(int) U FIRST(float) = {char, int, float}
- FIRST(<lista\_ids>) ::= FIRST(ID<lista\_ids'>) = {ID}dista\_ids> ::= ID<lista\_ids'>
- FIRST( $\langle \text{lista ids'} \rangle ::= \text{FIRST}(,\text{ID} \langle \text{lista ids'} \rangle) \cup \{\epsilon\} = \{1, \epsilon\}$
- FIRST(<sequencia\_comandos>) ::= FIRST(<comando> <sequencia\_comandos>) U
   {ε} = FIRST(<comando>) U {ε} = {if, while, repeat, ID}
- FIRST(<comando>) ::= FIRST(<comando\_selecao>) U FIRST(
   <comando\_repeticao>) U FIRST(<comando\_atribuicao>) = {if, while, repeat, ID}
- FIRST(<comando\_selecao>) ::= FIRST(if ( <condicao> ) then <comando\_ou\_bloco> <comando\_selecao'>) = {if}
- FIRST(<comando seleção'>) ::= FIRST(else <comando ou bloçõ>) U  $\{\epsilon\}$  =  $\{$ else $,\epsilon\}$

- FIRST(<comando\_ou\_bloco>) ::= FIRST(<comando>) U FIRST(<bloco>) = {if, while, repeat, ID, begin}
  - o **<comando ou bloco> ::= <**comando> | <bloco>
- FIRST(<comando\_repeticao>) ::= FIRST(while (<condicao> ) do
   <comando\_ou\_bloco>) U FIRST(repeat <comando\_ou\_bloco> until (<condicao> );)
   = {while,repeat}
- FIRST(<comando atribuicao>) ::= FIRST(ID := <expressao> ;) = {ID}
- FIRST(<condicao>) ::= FIRST(<expressão> relop <expressão>) = FIRST(<expressão>) = {num, caractere, ID, (}
  - **<condicao> ::=** <expressão> **relop** <expressão>
- FIRST(<expressao>) ::=FIRST( <termo><expressao'>) = FIRST(<termo>) = {num, caractere, ID, (}
  - o <expressao> ::= <termo><expressao'>
- FIRST(<expressao'>) ::= FIRST(+ <termo> <expressao'>) U FIRST( <termo> <expressao'>) U {ε} = {+, -, ε}
- FIRST(<termo>) ::= FIRST(<exp><termo'>) = FIRST(<exp>) = {num, caractere, ID, (}
- FIRST(<termo'>) ::= FIRST(\* <exp> <termo'>) U FIRST(/ <exp> <termo'>) U { $\epsilon$ } = {\*, /,  $\epsilon$ }
- FIRST(<exp>) ::= FIRST(<fator><exp'>) = FIRST(<fator>) = {num, caractere, ID, (}
- FIRST( $\langle \exp' \rangle$ ) ::= FIRST(\*\*  $\langle \text{fator} \rangle \langle \exp' \rangle$ ) U { $\epsilon$ } = {\*\*,  $\epsilon$ }
- FIRST(<fator>) ::= FIRST(num) U FIRST(caractere) U FIRST(ID) U FIRST((<expressao>)) = {num, caractere, ID, (}

Símbolos (Terminais e Não-Terminais)	FIRST
main	main
ID	ID
begin	begin
end	end

int	int
char	char
float	float
:	:
(	(
)	)
• • • • • • • • • • • • • • • • • • • •	·
if	if
then	then
else	else
while	while
do	do
repeat	repeat
until	until
relop	relop
:=	:=
,	,
+	+
-	-
*	*
**	**
/	/
num	num
caractere	caractere
<fator></fator>	{num, ID, (, caractere}
<exp'></exp'>	{**, ε}
<exp></exp>	{num, ID, (, caractere}

<termo'></termo'>	{*, /, ε}
<termo></termo>	{num, ID, (, caractere}
<expressao'></expressao'>	{+, -, ε}
<expressao></expressao>	{num, ID, (, caractere}
<condicao></condicao>	{num, ID, (, caractere}
<comando_atribuicao></comando_atribuicao>	{ID}
<comando_repeticao></comando_repeticao>	{while, repeat}
<comando_selecao'></comando_selecao'>	{else, ε }
<comando_selecao></comando_selecao>	{if}
<comando></comando>	{while, repeat, if, ID}
<blood></blood>	{begin}
<comando_ou_bloco></comando_ou_bloco>	{while, repeat, if, ID, begin}
<sequencia comandos="" de=""></sequencia>	{while, repeat, if, ID, ε}
<li>lista_ids'&gt;</li>	{, , ε}
<li>lista_ids&gt;</li>	{ID}
<tipo></tipo>	{char, int, float}
<declaracao_variavel></declaracao_variavel>	{char, int, float}
<declaracao_variaveis></declaracao_variaveis>	{char, int, float, ε}
<inicio></inicio>	{main}

## 3.3 Cálculo e Tabela de FOLLOW

- FOLLOW(<INICIO>) ::= FIRST(\$) = {\$} • **<INICIO> ::= main ID** <bloco>
- FOLLOW(<bloco>) ::= FOLLOW(<comando\_ou\_bloco>) U FOLLOW(<bloco>) = {else,while, repeat, if, ID, end, until}
- FOLLOW(<declaracao\_variaveis>) ::= FIRST(<sequencia\_comandos>end) {ε} U FOLLOW(<declaracao\_variaveis>) = {while,repeat,if,ID,end}

- FOLLOW(<declaracao\_variavel>) ::= FIRST(<declaracao\_variaveis>) {ε} U FOLLOW(<declaracao\_variaveis) = {char, int, float, while, repeat, if, ID, end}
- FOLLOW(<tipo>) ::= FIRST( : ids> ;) = {:}
- FOLLOW(<lista ids>) ::= FIRST(;) = {;}
- FOLLOW(<lista ids'>) ::= FOLLOW(<lista ids>) U FOLLOW(<lista ids'>) = {;}
- FOLLOW(<sequencia\_comandos>) ::= FIRST(end) U
   FOLLOW(<sequencia\_comandos>) = {end}
  - **<sequencia comandos>** ::= <comando> <sequencia comandos> | ε
- FOLLOW(<comando>) ::= FIRST(<sequencia\_comandos>) {ε} U FOLLOW(<comando\_ou\_bloco) = {while, repeat, if, ID, end, else, until}

  - <sequencia comandos>::= <comando> <sequencia comandos> | ε
  - o **<comando ou bloco> ::= <**comando> | <bloco>
- FOLLOW(<comando\_selecao>) ::= FOLLOW(<comando>) = {while, repeat, if, ID, end}
- FOLLOW(<comando\_selecao'>) ::= FOLLOW(<comando\_selecao>) = {while, repeat, if, ID, end}
- FOLLOW(<comando\_ou\_bloco>) ::= FIRST(<comando\_selecao'>) {ε} U
  FOLLOW(<comando\_selecao'>) U FOLLOW(<comando\_repeticao>) U FIRST(until
  ( <condicao> );) = {while, repeat, if, ID, end, else, until}
- FOLLOW(<comando\_repeticao>) ::= FOLLOW(<comando>) = {while, repeat, if, ID, end}
- FOLLOW(<comando\_atribuicao>) ::= FOLLOW(<comando>) = {while, repeat, if, ID, end}
- FOLLOW(<condicao>) ::= FIRST( ) then <comando\_ou\_bloco>) U FIRST( ) do <comando ou bloco>) U FIRST( ); ) = {)}
  - <condicao> ::= <expressão> relop <expressão>
  - comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco>
    <comando\_selecao'>

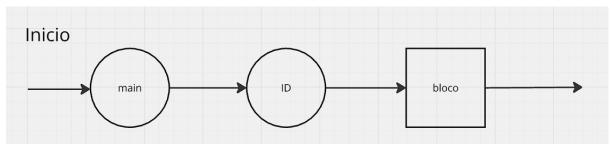
- comando\_repeticao> ::= while ( <condicao> ) do <comando\_ou\_bloco> |
  repeat <comando\_ou\_bloco> until ( <condicao> );
- FOLLOW(<expressao>) ::= FIRST(;) U FIRST(relop <expressão>) U FIRST()) = {; , relop, )}
  - o <expressao> ::= <termo><expressao'>
  - o <comando atribuicao> ::= ID := <expressao> ;
  - condicao> ::= <expressao> relop <expressao>
  - < <fator> ::= num | caractere | ID | (<expressao>)
- FOLLOW(<expressao'>) ::= FOLLOW(<expressao'>) U FOLLOW(<expressao'>) = {; , relop, }}
- FOLLOW( $\langle \text{termo} \rangle$ ) ::= FIRST( $\langle \text{expressao'} \rangle$ ) { $\epsilon$ } = {+, -}
  - o <termo> ::= <exp><termo'>
  - o <expressao> ::= <termo><expressao'>
  - **<expressao'>** := + <termo> <expressao'> | <termo> <expressao'> | ε
- FOLLOW(<termo'>) ::= FOLLOW(<termo>) U FOLLOW(<termo'>) = {+, -}
- FOLLOW( $\langle \exp \rangle$ ) ::= FIRST( $\langle \text{termo'} \rangle$ ) { $\varepsilon$ } = {\*, /}
- FOLLOW(<exp'>) ::= FOLLOW(<exp>) U FOLLOW(<exp'>) = {\*,/}
- FOLLOW( $\langle \text{fator} \rangle ::= \text{FIRST}(\langle \text{exp'} \rangle) \{\epsilon\} = \{**\}$

Símbolos Não-Terminais	FOLLOW
<inicio></inicio>	<b>{\$</b> }
<blook></blook>	{while, repeat, if, ID, end, else, \$, until}
<declaracao_variaveis></declaracao_variaveis>	{while, repeat, if, ID, end}
<declaracao_variavel></declaracao_variavel>	{char, int, float, while, repeat, if, ID, end}
<tipo></tipo>	{:}
<li>lista_ids&gt;</li>	{;}
<li>lista_ids'&gt;</li>	{;}
<sequencia_comandos></sequencia_comandos>	{end}
<comando></comando>	{while, repeat, if, ID, end, else, until}

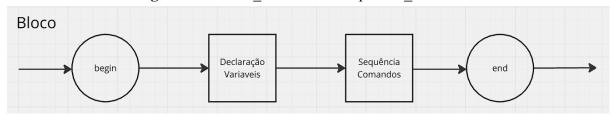
<comando_ou_bloco></comando_ou_bloco>	{while, repeat, if, ID, end, else, until}
<comando_selecao></comando_selecao>	{while, repeat, if, ID, end, else, until}
<comando_selecao'></comando_selecao'>	{while, repeat, if, ID, end, else, until}
<comando_atribuicao></comando_atribuicao>	{while, repeat, if, ID, end, else, until}
<comando_repeticao></comando_repeticao>	{while, repeat, if, ID, end, else, until}
<condicao></condicao>	{)}
<expressao></expressao>	{;, ), relop}
<expressao'></expressao'>	{;, ), relop}
<termo></termo>	{+, -, ;, ), relop}
<termo'></termo'>	{+, -, ;, ), relop}
<exp></exp>	{*, / , +, -, ;, ), relop}
<exp'></exp'>	{*, / , +, -, ;, ), relop}
<fator></fator>	{**, *, / , +, -, ;, ), relop}

# 3.4 Construção dos Grafos Sintáticos

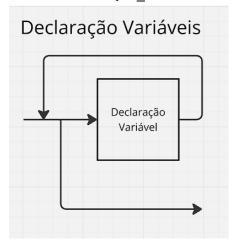
• <INICIO> ::= main ID <bloco>



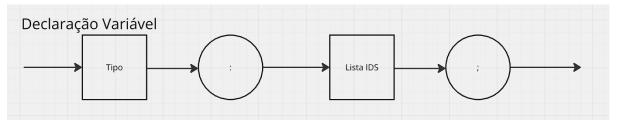
• **<blood> ::= begin <**declaracao\_variaveis> <sequencia\_comandos> **end** 



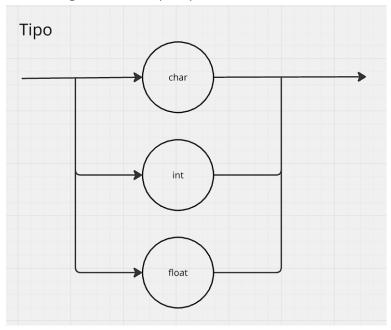
• <declaração\_variaveis> ::= <declaracao\_variavel> <declaracao\_variaveis> | ε



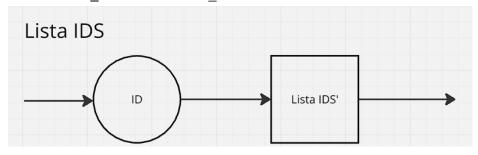
• <declaração\_variavel> ::= <tipo> : !:= <tipo> : !:= <tipo> :



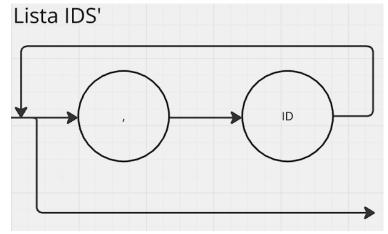
• <tipo> ::= char | int | float



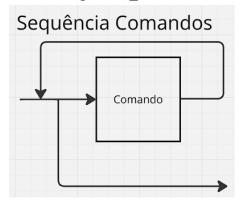
• lista\_ids> ::= ID<lista\_ids'>



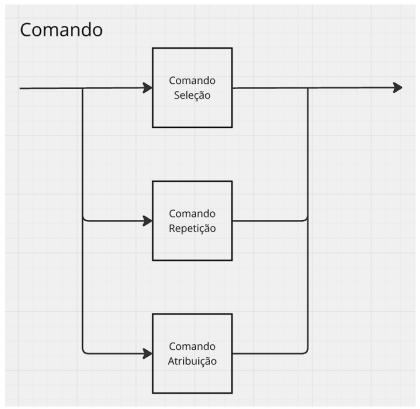
• - | Sista\_ids'> ::= ,ID



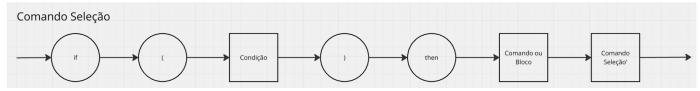
 $\bullet \quad \textbf{<sequencia\_comandos> ::=} < \hspace{-0.5em} \texttt{-comando>} < \hspace{-0.5em} \texttt{-sequencia\_comandos>} \mid \epsilon$ 



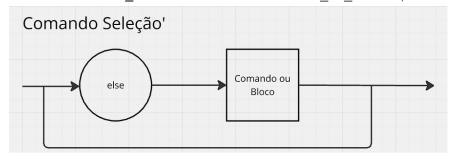
• <comando> ::= <comando\_selecao> | <comando\_repeticao> | <comando\_atribuicao>



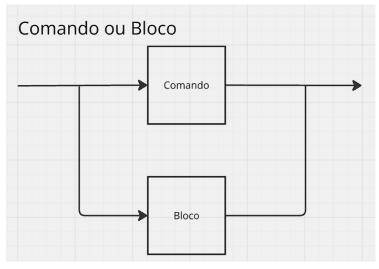
<comando\_selecao> ::= if ( <condicao> ) then <comando\_ou\_bloco>
 <comando\_selecao'>



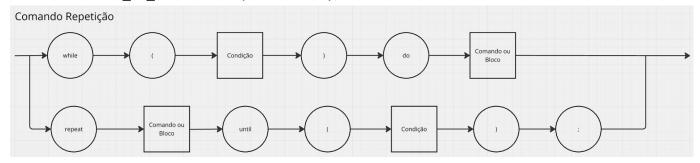
• <comando\_selecao'> ::= else <comando\_ou\_bloco> |  $\epsilon$ 



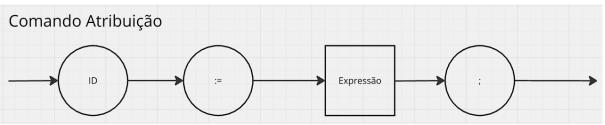
• <comando\_ou\_bloco> ::= <comando> | <bloco>



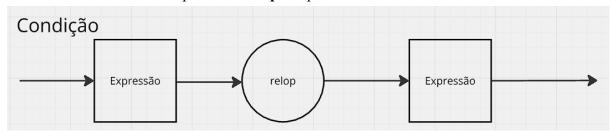
• <comando\_repeticao> ::= while ( <condicao> ) do <comando\_ou\_bloco> | repeat <comando\_ou\_bloco> until ( <condicao> );



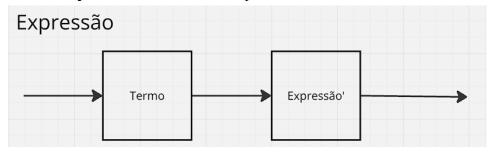
• <comando\_atribuicao> ::= ID := <expressao> ;



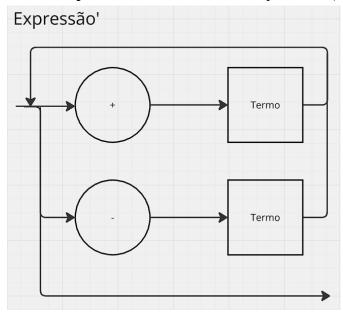
• **<condicao>** ::= <expressão> **relop** <expressão>



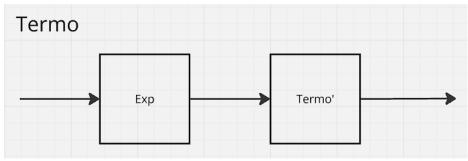
• <expressao> ::= <termo><expressao'>



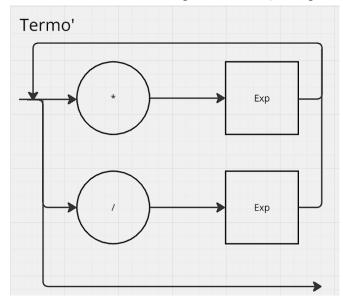
• **<expressao'>** ::= + <termo> <expressao'> | - <termo> <expressao'> | ε



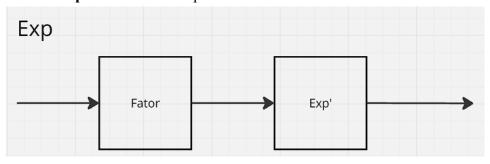
• <termo> ::= <exp><termo'>



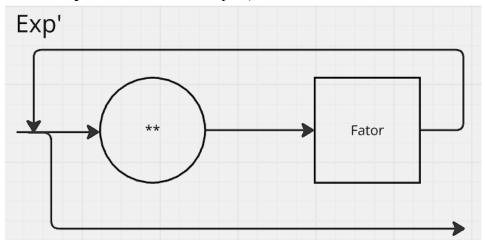
• <termo'> ::= \* <exp> <termo'> | / <exp> <termo'> |  $\epsilon$ 



• <exp>::= <fator><exp'>



•  $\langle exp' \rangle ::= ** \langle fator \rangle \langle exp' \rangle \mid \epsilon$ 



• <fator> ::= num | caractere | ID | (<expressao>)

