

Trabalho compiladores – GBC071 (2024-2)

Projeto da Linguagem

1. Definição da gramática livre de contexto (GLC):

a. $G(V, \Sigma, P, S)$:

- V = Conjunto dos símbolos não terminais;
- Σ = Conjunto dos símbolos terminais;
- P = Regras de produções;
- S = Símbolo/Estado inicial

b. Conjunto dos símbolos terminais (Σ) ou **tokens**:

- programa
- ID
- (
-)
- {%
- %}
- int
- char
- float
- :
- ;
- ,
- if
- [
-]
- then
- elseif
- else
- COMMENT
- while
- do
- :=
- relop
- +
- -
- *
- /
- **
- caractere
- num
- ws

- EOF
- c. Conjunto dos símbolos não terminais (V):
- APLICATIVO
 - BLOCO
 - DEC_VARIAVEIS
 - DEC_VARIAVEL
 - TIPO
 - LISTA_IDS
 - LISTA_ID
 - SEQ_CMD
 - COMAND
 - CMD_ATRIBUICAO
 - BLOCO_CMD
 - CMD_IF
 - CMD_IF_OPCIONAL
 - CMD_WHILE
 - CMD_DO_WHILE
 - CONDICAO
 - EXPRESSAO
- d. Regras de Produções (P) do conjunto V:
- APLICATIVO \rightarrow "programa" "ID" "(" ")" BLOCO
 - BLOCO \rightarrow "{" DEC_VARIAVEIS SEQ_CMD "%}"
 - DEC_VARIAVEIS \rightarrow DEC_VARIAVEL DEC_VARIAVEIS
 - DEC_VARIAVEIS $\rightarrow \varepsilon$
 - DEC_VARIAVEL \rightarrow TIPO ":" LISTA_IDS ";"
 - TIPO \rightarrow "int"
 - TIPO \rightarrow "float"
 - TIPO \rightarrow "char"
 - LISTA_IDS \rightarrow "ID" LISTA_IDS'
 - LISTA_IDS' \rightarrow ";" "ID" LISTA_IDS'
 - LISTA_IDS' $\rightarrow \varepsilon$
 - SEQ_CMD \rightarrow COMAND SEQ_CMD
 - SEQ_CMD $\rightarrow \varepsilon$
 - COMAND \rightarrow CMD_IF
 - COMAND \rightarrow CMD_WHILE
 - COMAND \rightarrow CMD_DO_WHILE
 - COMAND \rightarrow CMD_ATRIBUICAO
 - CMD_ATRIBUICAO \rightarrow "ID" ":"=" EXPRE
 - CMD_IF \rightarrow "if" "[" CONDICAO "]" "then" CMD_OR_BLOCO
CMD_IF_TAIL
 - CMD_IF_TAIL \rightarrow CMD_ELSEIF_LIST CMD_ELSE_OPT
 - CMD_ELSEIF_LIST \rightarrow "elseif" "[" CONDICAO "]" "then"
CMD_OR_BLOCO CMD_ELSEIF_LIST
 - CMD_ELSEIF_LIST $\rightarrow \varepsilon$
 - CMD_ELSE_OPT \rightarrow "else" CMD_OR_BLOCO
 - CMD_ELSE_OPT $\rightarrow \varepsilon$

- $CMD_WHILE \rightarrow "while" "[" CONDICAO "]" CMD_OR_BLOCO$
- $CMD_DO_WHILE \rightarrow "do" CMD_OR_BLOCO "while" "[" CONDICAO "]"$
- $CMD_OR_BLOCO \rightarrow COMAND$
- $CMD_OR_BLOCO \rightarrow BLOCO$
- $CONDICAO \rightarrow EXPRE "relop" EXPRE$
- $EXPRE \rightarrow TERMO EXPRE'$
- $EXPRE' \rightarrow "+" TERMO EXPRE'$
- $EXPRE' \rightarrow "-" TERMO EXPRE'$
- $EXPRE' \rightarrow \varepsilon$
- $TERMO \rightarrow EXP TERMO'$
- $TERMO' \rightarrow "*" EXP TERMO'$
- $TERMO' \rightarrow "/" EXP TERMO'$
- $TERMO' \rightarrow \varepsilon$
- $EXP \rightarrow FATOR EXP'$
- $EXP' \rightarrow "**" FATOR EXP'$
- $EXP' \rightarrow \varepsilon$
- $FATOR \rightarrow "-" FATOR$
- $FATOR \rightarrow "caractere"$
- $FATOR \rightarrow "num"$
- $FATOR \rightarrow "ID"$
- $FATOR \rightarrow "(" EXPRE ")"$

e. Expressões Regulares (dos tokens):

- $:$ \rightarrow $:$
- $;$ \rightarrow $;$
- $,$ \rightarrow $,$
- $($ \rightarrow $($
- $)$ \rightarrow $)$
- $\{$ \rightarrow $\{$
- $\}$ \rightarrow $\}$
- $:=$ \rightarrow $:=$
- $+$ \rightarrow $\backslash +$
- $-$ \rightarrow $-$
- $*$ \rightarrow $\backslash *$
- $/$ \rightarrow $/$
- $**$ \rightarrow $\backslash * \backslash *$
- $letra \rightarrow [A-Z a-z]$
- $digito \rightarrow [0-9]$
- $digitos \rightarrow digito +$
- $ID \rightarrow [letra _][letra digito _]*$
- $relop \rightarrow < | > | <= | >= | = | <>$
- $ws \rightarrow (" " | \backslash t | \backslash n | \backslash r)$
- $caractere \rightarrow ['^']$
- $num \rightarrow digitos (\backslash . digitos)? (E[\+-]? digitos)?$
- $COMMENT \rightarrow \{ \# [^ \#] \}^* \#$

f. Tabela de identificação dos tokens:

Identificação dos Tokens:

lexema	Nome do Token	Valor do Atributo
program	PROGRAM_TK	-
if	IF_TK	-
then	THEN_TK	-
else	ELSE_TK	-
Elseif	ELSEIF_TK	-
while	WHILE_TK	-
do	DO_TK	-
{# ... #}	COMMENT_TK	-
{% e %}	BLOCK_TK	-
:=	TOK_DOISPONTE Q	-
=	RELOP	EQ
<>	RELOP	NE
<	RELOP	LT
>	RELOP	GT
<=	RELOP	LE
>=	RELOP	GE
+	PLUS_TK	-
-	MINUS_TK	-
*	MULT_TK	-
/	DIV_TK	-
**	EXP_TK	-
int	TYPE_INT_TK	-
char	TYPE_CHAR_TK	-
float	TYPE_FLOAT_TK	-

:	TOK_DOISPONT	-
;	SEMICOLON_TK	-
,	COMMA_TK	-
[LBRACKET_TK	-
]	RBRACKET_TK	-
Identificadores	TOK_ID	Posição na tabela de símbolos
Qualquer número	TOK_NUM	Posição na tabela de símbolos
Caracteres	TOK_CARAC	Posição na tabela de símbolos

First

- $\text{FIRST}(\text{APLICATIVO}) = \{\text{"programa"}\}$
- $\text{FIRST}(\text{BLOCO}) = \{\text{"{"}\}$
- $\text{FIRST}(\text{TIPO}) = \{\text{"int", "float", "char"}\}$
- $\text{FIRST}(\text{DEC_VARIABEL}) = \text{FIRST}(\text{TIPO}) = \{\text{"int", "float", "char"}\}$
- $\text{FIRST}(\text{DEC_VARIAVEIS}) = \{\text{"int", "float", "char"}\} \cup \{\epsilon\}$
- $\text{FIRST}(\text{LISTA_IDS}) = \{\text{"ID"}\}$
- $\text{FIRST}(\text{LISTA_IDS'}) = \{\text{"", "epsilon"}\}$
- $\text{FIRST}(\text{SEQ_CMD}) = \{\text{"ID", "if", "while", "do"}\}$
- $\text{FIRST}(\text{CMD_OR_BLOCO}) = \{\text{"ID", "if", "while", "do", "{"}\}$
- $\text{FIRST}(\text{COMAND}) = \{\text{"ID", "if", "while", "do"}\}$
- $\text{FIRST}(\text{CMD_ATRIBUICAO}) = \{\text{"ID"}\}$
- $\text{FIRST}(\text{CMD_IF}) = \{\text{"if"}\}$
- $\text{FIRST}(\text{CMD_IF_TAIL}) = \{\text{"elseif", "else"}\} \cup \{\epsilon\}$
- $\text{FIRST}(\text{CMD_ELSEIF_LIST}) = \{\text{"elseif"}\} \cup \{\epsilon\}$
- $\text{FIRST}(\text{CMD_ELSE_OPT}) = \{\text{"else"}\} \cup \{\epsilon\}$
- $\text{FIRST}(\text{CMD_WHILE}) = \{\text{"while"}\}$
- $\text{FIRST}(\text{CMD_DO_WHILE}) = \{\text{"do"}\}$
- $\text{FIRST}(\text{CONDICAO}) = \{\text{"ID", "(", "num", "caractere"}\}$
- $\text{FIRST}(\text{EXPRE}) = \{\text{"ID", "(", "num", "caractere"}\}$
- $\text{FIRST}(\text{EXPRE'}) = \{\text{"relop", "+", "-"}\} \cup \{\epsilon\}$
- $\text{FIRST}(\text{TERMO}) = \{\text{"ID", "(", "num", "caractere"}\}$
- $\text{FIRST}(\text{TERMO'}) = \{\text{"*", "/"}\} \cup \{\epsilon\}$

- $FIRST(EXP) = \{ "ID", "(", "num", "caractere" \}$
- $FIRST(EXP') = \{ "+", "-", "*" \} \cup \{ \epsilon \}$
- $FIRST(FATOR) = \{ "ID", "(", "num", "caractere" \}$

Follow

- $FOLLOW(APLICATIVO) = \{ \$ \}$
- $FOLLOW(BLOCO) = \{ "\%" \}$
- $FOLLOW(DEC_VARIAVEIS) = \{ "if", "while", "do", "ID", "\%" \}$
- $FOLLOW(DEC_VARIAVEL) = \{ ":", ";" \}$
- $FOLLOW(TIPO) = \{ ":", ";" \}$
- $FOLLOW(LISTA_IDS) = \{ ":", ";" \}$
- $FOLLOW(LISTA_IDS) = \{ ":", ";" \}$
- $FOLLOW(SEQ_CMD) = \{ "\%", \$ \}$
- $FOLLOW(COMAND) = \{ "else", "elseif", "\%", \$ \}$
- $FOLLOW(CMD_ATRIBUICAO) = \{ ":", ";" \}$
- $FOLLOW(BLOCO_CMD) = \{ "else", "elseif", "\%", \$ \}$
- $FOLLOW(CMD_IF) = \{ "\%", \$ \}$
- $FOLLOW(CMD_IF_TAIL) = \{ "\%", \$ \}$
- $FOLLOW(CMD_ELSEIF_LIST) = \{ "else", "\%", \$ \}$
- $FOLLOW(CMD_ELSEIF) = \{ "elseif", "else", "\%", \$ \}$
- $FOLLOW(CMD_ELSE_OPT) = \{ "\%", \$ \}$
- $FOLLOW(CMD_WHILE) = \{ "\%", \$ \}$
- $FOLLOW(CMD_DO_WHILE) = \{ "\%", \$ \}$
- $FOLLOW(CONDICAO) = \{ "]" \}$
- $FOLLOW(EXPRE) = \{ "relop", ")", "]" \}$
- $FOLLOW(EXPRE) = \{ "relop", ")", "]" \}$
- $FOLLOW(TERMO) = \{ "+", "-", "relop", ")", "]" \}$
- $FOLLOW(TERMO) = \{ "+", "-", "relop", ")", "]" \}$
- $FOLLOW(EXP) = \{ "+", "-", "*", "/", "relop", ")", "]" \}$
- $FOLLOW(EXP) = \{ "+", "-", "*", "/", "relop", ")", "]" \}$
- $FOLLOW(FATOR) = \{ "+", "-", "*", "/", "relop", ")", "]" \}$