

## Especificação da linguagem

Seguem algumas definições que serão utilizadas posteriormente:

- $D = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
- $L = \{A, B, \dots, Z, a, \dots, z\}$
- $S = \{\sim, \text{`}, !, @, \#, \$, \%, \wedge, \&, *, (, ), \_, -, +, \backslash, /, [, ], |, \backslash, ?, >, <, ., , \}$

A tabela a seguir descreve os lexemas que devem ser identificados para geração de *tokens* pelo analisador léxico para a linguagem Irrisória.

Token	Descrição informal	Padrão
numero	Constante numérica (inteira ou real)	$D+(\backslash .D+)?((E e)(+ -)?D+)?$
literal	Constante literal	$"(L D S \backslash t)^{*}"$
identificador	Nome de variável	$L(L D _)^{*}$
op_relacional	Operador relacional	$< > >= <= == <>$
op_aritmetico	Operador aritmético	$+ - * /$
atribuicao	Atribuição	$=$
abre_par	Abre parênteses	$($
fecha_par	Fecha parênteses	$)$
virgula	Vírgula	$,$
ponto_e_virg	Ponto e vírgula	$;$
vazio	Espaço, tabulação e salto de linha	$( \backslash t \backslash r \backslash n)^{+}$
comentario	Comentários do programador que podem ser retornados como token ou descartados	$\{(L D S \backslash t)^{*}\}$
fim_de_arq	Final do arquivo	Informado pelo sistema operacional ao encerrar a leitura do arquivo com o código fonte
erro	Qualquer símbolo que não pertença ao alfabeto da linguagem	

A seguir, são apresentadas as palavras reservadas da linguagem Irrisória e que, portanto, também devem ser reconhecidas e ter seus respectivos *tokens* retornados. Nesse caso, o *token* pode coincidir com a palavra reservada.

- **inicioprog** – Delimita o início do programa.
- **fimprog** – Delimita o fim do programa.
- **inicioargs** – Delimita o início dos argumentos passados como entrada para o programa.
- **fimargs** – Delimita o fim dos argumentos passados como entrada para o programa.
- **iniciovars** – Delimita o início da declaração de variáveis.
- **fimvars** – Delimita o fim da declaração de variáveis.
- **escreva** – Escreve na saída padrão.
- **inteiro** – Tipo de dado inteiro.
- **real** – Tipo de dado real.
- **literal** – Tipo de dado literal.
- **se** – Elemento de estrutura condicional.
- **entao** – Elemento de estrutura condicional.
- **fimse** – Elemento de estrutura condicional.
- **enquanto** – Elemento de estrutura de repetição.
- **faca** – Elemento de estrutura de repetição.
- **fimenquanto** – Elemento de estrutura de repetição.