



Departamento de
Informática

2º Ciclo Engenharia Informática
Linguagens de Programação e Compiladores
2023/2024

Nome:

LabTP#02 | 19/10/2023

Número de Aluno:

Duração: 60 min. + 5 min. tolerância

Tópicos

- *Análise lexical com a ferramenta Flex.*

Exercício 2

Identificação de literais em Pascal

Realize um programa que identifique os seguintes elementos de um programa em Pascal: identificadores, números inteiros, números reais em vírgula flutuante de precisão simples e cadeias de caracteres. Ignore a possibilidade de a cadeia de caracteres conter aspas (").

Na realização do exercício, assuma as seguintes definições feitas pelo YACC (nesta fase, apenas interessa saber que as definições correspondem a tokens e que a %union serve para comunicar dados entre o Flex e o YACC):

```
%union int i; float r; char *s;
```

```
%token<s> ID STR
```

```
%token<i> INT
```

```
%token<r> REAL
```

```
%token ASSIGN LE GE NE DIV MOD AND NIL NOT BEGIn END
```

```
%token IF THEN ELSE WHILE DO REPEAT UNTIL FOR TO DOWNT0
```

```
%token PROGRAM LABEL CONST TYPE VAR PROCEDURE FUNCTION
```

```
%token PACKED ARRAY RECORD OF SET FILE CASE WITH GOTO
```

Note-se que os nomes dos tokens correspondem a macros em C/C++ e que devem ser escolhidos por forma a não colidirem com outras definições (e.g., nomes de classes).

Exercício 2

Realize um programa que permita extrair, imprimindo no terminal, as cadeias de caracteres em programas nas linguagens: Smalltalk, C e PostScript.

Em Smalltalk, uma cadeia de caracteres é delimitada pelo carácter ', podendo este carácter ser introduzido na cadeia através da sua duplicação: 'batata"cebola'.

Em C, uma cadeia de caracteres é delimitada pelo carácter aspa ", podendo este carácter ser introduzido na cadeia através da sequência (\"): "batata\"cebola".

Em PostScript, uma cadeia de caracteres é delimitada por parenteses curvos, podendo os parenteses ser introduzidos livremente no texto se estiverem emparelhados. Parenteses não emparelhados podem também ser introduzidos se precedidos do carácter (\): (batata)cebola).