**TRABALHO nO1: especificação dos *tokens* da linguagem 2020.2**

**INSTRUÇÕES:**

1. O trabalho deve ser desenvolvido pela mesma equipe do desenvolvimento do compilador.
2. A especificação do TRABALHO nº1 estará disponível no AVA3, na aba COMPILADOR, arquivo trabalho no1.doc, a partir das **22h30 do dia 31 de agosto**.
3. Coloque os nomes dos integrantes da equipe e as respostas das questões nos campos apropriados.
4. Poste o arquivo com as respostas na **pasta** da sua equipe no AVA3 até às **9h do dia 08 de setembro**.

**EQUIPE (a mesma do desenvolvimento do compilador):**

|  |
| --- |
| **Jadiel dos Santos**  **Gian Carlo Giovanella**  **Matheus Mahnke** |

Para as linguagens regulares (dos tokens) denotadas abaixo, especifique **definições regulares**:

1. **(2,0 pontos) identificador**:é qualquer palavra composta por letras minúsculas, dígitos e *underline* ( \_ ). Um identificador deve começar com uma letra minúscula. Na sequência podem existir outras ocorrências (zero ou mais) de letras minúscula ou de dígitos ou de *underlines*, desde que não possua *underlines* consecutivos e não termine com *underline*.
2. **(1,5 pontos)** **constante** int:é qualquer palavra composta por no mínimo um dígito decimal (0 a 9).
3. **(1,5 pontos)** **constante** float:é qualquer palavra composta por uma parte inteira, concatenada com uma parte fracionária, nessa ordem. A parte inteira é composta por no mínimo um dígito decimal. A parte fracionária é composta por um ponto decimal ( **\.** ), seguido por no mínimo um dígito decimal.
4. **(2,0 pontos)** **constante** str: pode ser escrita de uma das seguintes formas:

d.1) é qualquer palavra delimitada por aspas duplas ( **\"** ), uma no início e outra no final. Uma constante str pode possuir entre as aspas duplas zero ou mais ocorrências de quaisquer caracteres, exceto aspa dupla ( **\"** ) e quebra de linha ( **\n** );

d.2) é qualquer palavra delimitada por aspas simples ( **\'** ), uma no início e outra no final. Uma constante str pode possuir entre as aspas simples zero ou mais ocorrências de quaisquer caracteres, exceto aspa simples ( **\'** ), quebra de linha ( **\n** ) e barra invertida ( **\\** ).

1. **(1,0 ponto)** **comentário de linha**: é qualquer palavra que inicia com *hashtag* ( **\#** ) seguida por zero ou mais ocorrências de quaisquer caracteres, exceto quebra de linha ( **\n** ).
2. (**2,0 pontos**) **comentário de bloco**: é qualquer palavra que inicia com abre colchetes ( **\[** ) seguido de quebra de linha ( **\n** ) e termina com quebra de linha ( **\n** ) seguida de fecha colchetes ( **\]** ). Um comentário de bloco pode possuir entre os símbolos iniciais e finais zero ou mais ocorrências de quaisquer caracteres, exceto abre e fecha colchetes .

**colocar aqui as respostas:**

| **definições regulares AUXILIARES** (válidas para todas as especificações abaixo) |
| --- |
| **mai: [A-Z]**  **min: [a-z]**  **letra: mai | min**  **dig: [0-9]**  **inteiro: dig+**  **fracionario: \. inteiro** |

|  |
| --- |
| **definições regulares para os TOKENS** |
| **id: min ((min|dig)? \_? (min|dig))\***  **int: inteiro**  **float: inteiro fracionario**  **str\_simples: \' [^\'\n\\]\* \'**  **str\_dupla: \" [^\"\n]\* \"**  **comentario\_linha: \# [^\n]\***  **comentario\_bloco: \[ \n [^\[\]]\* \n \]** |