

Q1

- Ignorar espaços, tabulações, saltos de linha e comentários.
- Dividir os caracteres da entrada em lexemas e salvá-los em tokens para identificação posterior.
- Guardar linha e coluna para uma possível identificação de erro.

Q2

- Token é um objeto que armazena os lexemas. Contém nome e um atributo opcional.
Ex.: <+> <-> <num, 10>
- Padrão é uma descrição da forma que as sequências de caracteres podem assumir.
Ex.: a letra 'a', depois o número 54, depois '_', depois ';-' e por fim, qualquer letra minúscula.
- Lexema é uma sequência de caracteres no programa fonte que casa com o padrão para um token.

Q3

- Quando mais de um lexema casar com o padrão de um token.
Ex.: Os lexemas '5' e '6,0' casam com o padrão de um num. Os lexemas 'total', 'cont' e 'val' casam com o padrão de um 'id'.
- Quando for necessário no futuro do programa.

Q6

São identificadores de expressões regulares. Essa identificação pode ser usada em outras expressões regulares. Ex.: letra \rightarrow [a-zA-Z]. 'letra' é uma definição regular.