

Processamento de Linguagens

Ficha Introdutória para as Aulas Práticas sobre ER

Ano Lectivo de 20/21

Deteção de Tags HTML

Neste exercício, vais usar expressões regulares para detetar os elementos estruturais usadas numa página HTML.

Um elemento `elem` é composto por uma anotação (tag) de abertura `<elem>`, seguida de algum texto, seguido de uma tag de fecho `</elem>`. A barra de divisão na tag de fecho vem sempre antes do nome do elemento. A exceção a esta regra são as tags de elementos vazios. Estes elementos são constituídos apenas por uma tag com uma barra de divisão a seguir ao nome: `<elem/>`.

Eis alguns exemplos de elementos:

- O elemento `p` para os parágrafos: `<p>Isto é um parágrafo.</p>`
- Podem existir um ou mais espaços dentro da tag, antes ou depois do nome do elemento: `< p>Isto é um parágrafo.</p >`
- Um elemento vazio pode ser representado como apresentado em cima, `<p/>`, ou por duas tags sem qualquer conteúdo entre elas: `<p></p>`
- Alguns elementos também podem ter atributos, por exemplo, o elemento que serve para adicionar hiperlinks a uma página web: `Google`

Tarefa

Dada uma lista de linhas HTML, encontra o nome dos elementos (ignorando os atributos) e imprime-os numa única linha com os nomes ordenados alfabeticamente e separados por `','`: `elem1;elem2;elem3;...`

Formato de input

A primeira linha contém um inteiro, `N`, o número de linhas HTML que se seguem.

Restrições

- $1 \leq N \leq 100$
- Cada linha contém < 10000 caracteres
- As linhas de HTML foram retiradas da Wikipedia, alguma análise deste repositório poderá ajudar;
- Espaços iniciais e finais foram retirados das linhas HTML.

Formato de output

Escreve uma única linha com o nome de todos os elementos encontrados, sem repetições. Os nomes dos elementos deverão estar por ordem alfabética e separados por `','`.

Input exemplo

2

```
<p><a href="http://www.quackit.com/html/tutorial/html_links.cfm">Example Link</a></p>
<div class="more-info"><a href="http://www.qt.com/lexamples.cfm">More Link Examples...</a></div>
```

Output exemplo

```
a;div;p
```