

Volem **eliminar** certes paraules d'una seqüència de paraules a l'entrada, que representa una pàgina web. La seqüència està formada per paraules de dos tipus: paraules que són *tags* d'HTML, i paraules que no ho són. Del conjunt de les paraules que no són *tags* en parlarem com a "text".

Els *tags* HTML sempre tenen un nom delimitat per "<" i ">" i són de dos tipus: d'**obrir** (<nom>) i de **tancar** (</nom>). Noteu que els de tancar tenen una barra (/) com a segon caràcter. Es pot pensar que una parella de *tags* d'obrir i tancar són com dos "parèntesis" d'obrir i tancar, però **amb nom**, de tal manera que els podem distingir entre ells. Per exemple, l'equivalent a "<a>" seria "()", però en el cas dels *tags* podem distingir entre a i b.

Tornant a la seqüència, les paraules d'entrada formen una pàgina HTML correcta, que vol dir que tots els *tags* d'obrir tindran un *tag* de tancar corresponent, tal com passa amb els parèntesis en una expressió ben parentitzada, però aquest cop amb **correspondència de nom**. Per exemple, les seqüències següents, una per línia, són pàgines HTML correctes.

```
<html> a <head> b </head> <body> </body> c </html>
a <html> <head> b <t> c </t> </head> <body> <a> d </a> </body> </html>
<div> <p> a <span> </span> </p> b c <ul> <li> d </li> </ul> </div> e
```

També direm que en una seqüència com

```
<p> <f1> <x> </x> </f1> <f2> </f2> </p>
```

els *tags* <f1> i <f2> són **fills directes** del tag <p>. Això és perquè el tags <f1> i <f2> apareixen en l'àmbit delimitat per <p>, i no hi ha cap altre àmbit entremig. En aquest sentit, el tag <x> *no* és fill directe de <p> (ja que l'àmbit de <f1> està entremig de <p> i <x>), però <x> és fill directe de <f1>.

Així doncs, el text que volem eliminar és el que està **dins d'un tag <script>**, perquè podria ser codi JavaScript maliciós. Els *tags* <script> són especials en un fet important: a dins mai tenen altres *tags*, sempre tenen només text.

Ara bé, no podem esborrar tot el text a dins dels *tags* <script>, perquè hi ha una **excepció**: quan el *tag* <script> és fill directe de <head> i alhora aquest és fill directe d'<html>. Per exemple, a la seqüència

```
<html> <head> <script> keep! </script> </head> </html>
```

la paraula "keep!" no s'ha d'esborrar.

Codi disponible

Per poder fer el programa més còmodament, el codi públic d'aquest problema (la icona del gatet) proporciona:

- Un Makefile que compila el programa amb només fer make.
- Una carpeta .vscode que permet compilar i depurar només prement F5.

- Una classe `VStack` (`vstack.hh` i `vstack.cc`) amb dos mètodes especials: 1) **top**, que permet consultar el cim i també elements per sota del cim, i 2) **contains**, que permet saber si un element concret és a qualsevol posició de la pila.
- Funcions que treballen amb *tags* d'HTML (`html_elem.cc` i `html_elem.hh`).

Tots els mètodes i funcions del codi proporcionat estan documentats en el format Doxygen.

Entrada

Una seqüència de paraules (`strings`) amb paraules que són *tags* HTML i d'altres que no ho són. La seqüència forma una pàgina HTML correcta.

Sortida

La mateixa seqüència eliminant les paraules de dins dels *tags* `<script>`, exceptuant les paraules dins d'un `<script>` que és fill directe de `<head>`, i alhora aquest `<head>` és fill directe d'`<html>`. Després de cada paraula cal posar un sol espai.

Exemple d'entrada 1

```
<root> <html> <script> #$$@& </script> </html> </root>
```

Exemple de sortida 1

```
<root> <html> <script> </script> </html> </root>
```

Exemple d'entrada 2

```
<html>
  <script> voldemort </script>
  <head>
    <script> a b c d e </script>
  </head>
  <script> alert("pwned")! </script>
</html>
```

Exemple de sortida 2

```
<html> <script> </script> <head> <script> a b c d e </script>
```

Exemple d'entrada 3

```
e
<six> l
  <html> b
    <head> e
      <script> n
      </script> r
    </head> e
  </html> v
</six> q
```

Exemple de sortida 3

```
e <six> l <html> b <head> e <script> n </script> r </html>
```

Observació

Descarrega el codi públic (icona del gatet), i implementa `main.cc`. Envia només `main.cc` al Jutge, no tot el projecte.

Informació del problema

Autoria: Pau Fernández

Generació: 2026-03-24T11:36:43.063Z

© *Jutge.org*, 2006–2026.
<https://jutge.org>