

Java → Html Parser

Hard 19 minutes ? Share:

Code Challenge — Write a program

This exercise seems to be poorly explained. You can help us to improve it if you can solve it. Just write your description in the comments.

Today you will create your own HTML parser! You should get all the content from tags.

- For example: `<h4>
content</br></h4>`
- Answer: from `
`(*content*), after that from `<h4>`(`
content</br>`)

Note: all tags are valid.

Remember: The order of the contents on the same level of the hierarchy should be as in the input string.

Sample:

```
<html><a>hello</a><h1><h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6></h1><br>world</br></html>
```

Answer:

```
|hello
||nestedHello
||nestedWorld
|||top
|||<br>top</br>
|<h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6>
|world
|<a>hello</a><h1><h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6></h1><br>world</br>
```

Sample Input 1:

```
<html><a>hello</a><h1><h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6></h1><br>world</br></html>
```

Sample Output 1:

```
hello
nestedHello
nestedWorld
top
<br>top</br>
<h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6>
world
<a>hello</a><h1><h4>nestedHello</h4><h3>nestedWorld</h3><h6><br>top</br></h6></h1><br>world</br>
```

Sample Input 2:

```
<html><input><h2><a>hello</a><br>world</br></h2></input></html>
```

Sample Output 2:

```
hello
world
<a>hello</a><br>world</br>
<h2><a>hello</a><br>world</br></h2>
<input><h2><a>hello</a><br>world</br></h2></input>
```

Sample Input 3:

```
<h4><br>content</br></h4>
```

Sample Output 3:

```
content
<br>content</br>
```



If you don't see your IDE opened, switch to it manually

[Code snippets from theory](#)

[Show solution](#) [Skip problem](#)

Time limit: 8 seconds Memory limit: 256 MB

[Comments \(8\)](#)

[Hints \(0\)](#)

[Useful links \(0\)](#)

[Solutions \(0\)](#)

Share something, Sergey Kubatko

Post

Please do not post solutions here

Sort by: [Last posted](#) ▼

MK [Maciej Kuchta](#) [about 1 month ago](#) [Report](#)

The reference solution is simple and elegant. Mine was a lot more complicated, I couldn't see a way to solve it without regex, sadly. But it was a little eye-opening.

♡ 1 [Reply](#)

YV [Yuriy Volkovskiy \(β\)](#) [about 2 months ago](#) [Report](#)

Quite easy after Nesting Season.

♡ 0 [Reply](#)

DM [Dmitrij Morozov](#) [3 months ago](#) [Report](#)

You were required to write that the symbol / occurs only in tags. I learn Java not HTML.

♡ 1 [Reply](#)

S [SDWillems](#) [3 months ago](#) [Report](#)

I have NO idea what you want me to do here and what this has to do with Queue or Stack..

♡ 4 [Reply](#)

U1 [User 110292](#) [4 months ago](#) [Report](#)

Sorry, but I have to say that the explanation is very poor. I have no idea what to do here...

♡ 4 [Reply](#)

A [A.Ruskomaa](#) [5 months ago](#) [Report](#)

This exercise would really benefit from a better explanation.

♡ 1 [Reply](#)

FS [Full Stackly](#) [8 months ago](#) [Report](#)

wow, can anyone explain please the pattern of answer's order? I don't understand why is there `
top</br>` & `top` , but only `world` without `
world</br>`

♡ 3 [Reply](#)

MK **Maciej Kuchta** [about 1 month ago](#) [Report](#)

Because you search for contents enclosed by tags. `
world</br>` stands alone, not embraced by any.

♡ 0 [Reply](#)