

PRACTICE

COMPETE

JOBS

LEADERBOARD





Practice > Python > Regex and Parsing > HTML Parser - Part 1

HTML Parser - Part 1 ☆

84/115 challenges solved

Rank: **3074** | Points: **1635** (!)

Your HTML Parser - Part 1 submission got 30.00 points. Share

Try the next challenge | Try a Random Challenge

Tweet



Problem	Submissions	Leaderboard	Discussions	Editorial &
HTML				
Hypertext Markup	<i>Language</i> is a standard n	narkup language used f	or creating World Wide	e Web pages.
Parsing				
	ess of syntactic analysis ond describing their synta		involves resolving a str	ring into its
HTMLParser				
An <i>HTMLParser</i> ins	tance is fed HTML data a	nd calls handler metho	ds when start tags, end	d tags, text,
comments, and oth	er markup elements are	encountered.		
Example (based or	the original Python doc	umentation):		
Code				
from HTMLPars	er import HTMLParse	r		
def handlo print def handlo print def handlo	Parser(HTMLParser): e_starttag(self, tag "Found a start tag e_endtag(self, tag): "Found an end tag e_startendtag(self, "Found an empty tag	:", tag : ", tag tag, attrs):		
	the parser and fed	it some HTML		
	MLParser() ' <html><head><title: body><h1>HackerRank</h1></title: </head></html>			
Output				
Found a start Found a start Found a start Found an end Found a start Found a start Found an end Found an end Found an end Found an empt	tag : head tag : title tag : title tag : head tag : body tag : h1 tag : h1			

Author	DOSHI
Difficulty	Easy
Max Score	30
Submitted By	4034

NEED HELP?

- View discussions
- **W** View editorial
- View top submissions

RATE THIS CHALLENGE



MORE DETAILS

- Download problem statement
- Suggest Edits





```
Found an end tag : body
Found an end tag : html
```

.handle_starttag(tag, attrs)

This method is called to handle the *start tag* of an element. (For example: <div class='marks'>)

The *tag* argument is the name of the tag converted to lowercase.

The *attrs* argument is a list of (name, value) pairs containing the attributes found inside the tag's <> brackets.

.handle_endtag(tag)

This method is called to handle the end tag of an element. (For example: </div>)

The tag argument is the name of the tag converted to lowercase.

.handle_startendtag(tag,attrs)

This method is called to handle the *empty tag* of an element. (For example:

 />)

The tag argument is the name of the tag converted to lowercase.

The *attrs* argument is a list of (name, value) pairs containing the attributes found inside the tag's <>

Task

You are given an HTML code snippet of N lines.

Your task is to print start tags, end tags and empty tags separately.

Format your results in the following way:

```
Start : Tag1
End : Tag1
Start : Tag2
-> Attribute2[0] > Attribute_value2[0]
-> Attribute2[1] > Attribute_value2[1]
-> Attribute2[2] > Attribute_value2[2]
Start : Tag3
-> Attribute3[0] > None
Empty : Tag4
-> Attribute4[0] > Attribute_value4[0]
End : Tag3
End : Tag2
```

Here, the -> symbol indicates that the tag contains an attribute. It is immediately followed by the name of the attribute and the attribute value.

The > symbol acts as a separator of the attribute and the attribute value.

If an HTML tag has no attribute then simply print the name of the tag.

If an attribute has no attribute value then simply print the name of the attribute value as None.

Note: Do not detect any *HTML* tag, attribute or attribute value inside the *HTML* comment tags (<!-Comments -->).Comments can be multiline as well.

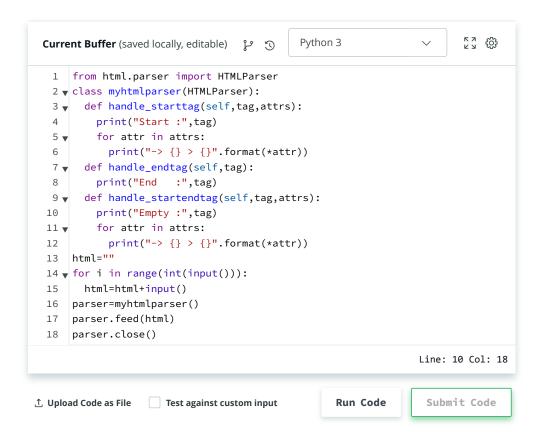
Input Format

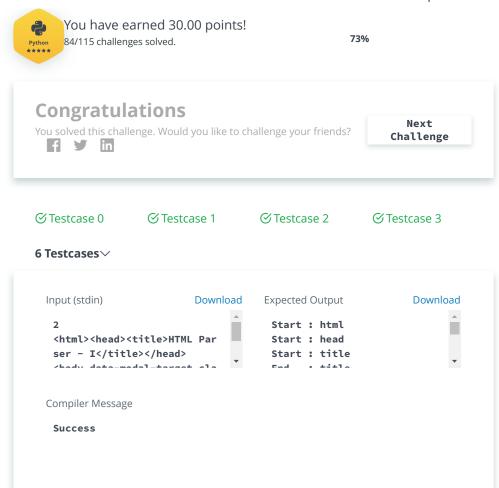
The first line contains integer N, the number of lines in a HTML code snippet.

The next N lines contain HTML code.

Constraints

```
• 0 < N < 100
Output Format
Print the HTML tags, attributes and attribute values in order of their occurrence from top to bottom in
the given snippet.
Use proper formatting as explained in the problem statement.
Sample Input
  <html><head><title>HTML Parser - I</title></head>
  <body data-modal-target class='1'><h1>HackerRank</h1><br /></body></html>
Sample Output
  Start: html
  Start: head
  Start: title
  End : title
  End : head
  Start : body
  -> data-modal-target > None
  -> class > 1
  Start: h1
  End: h1
  Empty : br
  End : body
  End : html
```





Contest Calendar | Blog | Scoring | Environment | FAQ | About Us | Support | Careers | Terms Of Service | Privacy Policy | Request a Feature