

microsoft.public.scripting.jscrip: Re: (repost): Very tough question, about HTML manipulation

Re: (repost): Very tough question, about HTML manipulation

Source: <http://www.tech-archive.net/Archive/Scripting/microsoft.public.scripting.jscrip/2004-07/0212.html>

From: Julian Turner (*JulianTurner_at_discussions.microsoft.com*)

Date: 07/13/04

Date: Tue, 13 Jul 2004 08:48:02 -0700

Hello Welman

I can see the point you are making.

You need to be able to apply a regular expression to chunks of text, so I can see why my individual character method would not work.

Nevertheless "Mission Impossible" just elicits inspiration, so I will try out some ideas on you, which may help you go forward.

I will not comment on your RegExps of course.

Important facts/assumptions:–

1. I am assuming beneath all of this that you want to work in IE. There is a different (and perhaps more helpful) DOM Range model for Mozilla and the like conforming to W3 standards. See www.w3.org. IE has its own unique model for Text Ranges.
2. You can get the `parentElement()` of a range. There may be a little complexity here for inter element boundaries, but I shall ignore this (apologies).
3. That it would be acceptable to you to apply the RegExp separately to the "Bit of text in bold" and "Bit of text in LI".

```
<P><B>Bold [SELECTION_BEGIN]Bit of text in bold</B></P>  
<UL><LI>Bit of text in LI [SELECTION_END]things</LI></UL>
```

This is the key issue really. If you want the RegExp to apply accross the whole selected text in one chunk, then it is certainly a more general problem, which I will not try to answer now.

If my more "special" case in 3 is acceptable to you, then I suggest the following.

The idea is to find sub-sets of the range that fall within different elements.

1. Again create a Cursor and Reference Range.

Re: (repost): Very tough question, about HTML manipulation

microsoft.public.scripting.jscript: Re: (repost): Very tough question, about HTML manipulation

> *W. Jordan*
>
> *"Julian Turner" <Julian.Turner@discussions.microsoft.com> wrote in message*
> *news:ADCB8363-FEF1-4FD1-861C-B2B4049D11F4@microsoft.com...*
> *: I have some thoughts for you, but this is not a complete answer. It is*
> *based on something I am working on myself along similar lines.*
> :
> *: I assume that all you want to do is change to upper case.*
> :
> *: Important facts:-*
> :
> *: - A TextRange can be as small as 1 character, or even nothing (i.e. a*
> *cursor insertion point).*
> :
> *: - You can have multiple text ranges over the same text.*
> :
> *: - setting just one character on the "text" property of the range, I*
> *believe would not affect the mark-up*
> :
> *: Suggested method*
> :
> *: 1. get the TextRange from the selection object. Let us call this the*
> *"Cursor", which we will use to go character by character.*
> :
> *: 2. Create a duplicate of the TextRange using the duplicate() method. You*
> *will use the duplicate to keep a "Reference" of the start and end points of*
> *the range.*
> :
> *: 3. Collapse the Cursor range, using the collapse() method, to the start*
> *of the Reference range. This collapses it so the start and end points are*
> *the same.*
> :
> *: 3. Expand the Cursor by one character, using the expand() method, and get*
> *the character, i.e. Cursor.text.*
> :
> *: 4. Is the character is an a-z character? E.g. in Regular Expression*
> *terms: /[a-z]/gi.test(Cursor.text).*
> :
> *: 5. If so, get the character, convert it to upper case (toUpperCase()),*
> *and save it back in Cursor.text. This is the key point, setting just one*
> *character on the text property, I believe would not affect the mark-up.*
> :
> *: 6. Collapse the Cursor to the end of the character, using collapse.*
> :
> *: 7. Compare the end point of the Cursor to the end point of the Reference*
> *using compareStartPoints method (I think that is what it is called) which*
> *allows you to compare the start and end points of two ranges.*
> :
> *: 8. If the Cursor is not at the end of the Reference, go to 3.*
> :
> *: It is cumbersome, but it is granularity should help.*

Re: (repost): Very tough question, about HTML manipulation

> :
> :
> :
> :
> : *"Welman Jordan" wrote:*
> :
> : > *Hello,*
> : >
> : > *I've posted this thread on IE.scripting for a month, but*
> : > *nobody answered me, thus I reposted it here, I hope*
> : > *somebody can help.*
> : >
> : >
> : > *I have got a very darn tough question that I can not solve*
> : > *by myself. Any JavaScript gurus please help!*
> : >
> : > *Here's the problem:*
> : >
> : > *I wanna replace some user selected text in a web page,*
> : > *however, after the replacement, the format of the selection*
> : > *should be preserved if possible.*
> : > -----
> : > *For example, here's a piece of code:*
> : >
> : > *<P>Bold [SELECTION_BEGIN]text</P>*
> : > *list item: item [SELECTION_END]things*
> : >
> : > *If it is rendered on a browser, it should looks like this way:*
> : > *a paragraph, bold text,*
> : > *and then a bulleted list, an item.*
> : >
> : > *The selection point begins in the middle of the bold text,*
> : > *and spans into the list item.*
> : > *If I copy the selected text, I will get:*
> : > *text*
> : > *list item: item*
> : >
> : > *Now, I wanna do something with the script, replace the*
> : > *selected text from "text", and "list item: item" to uppercase*
> : > *"TEXT", and "LIST ITEM: ITEM". However, the "text"*
> : > *should still be bold, and the "list item: item" should still*
> : > *stay in the bulleted list. How to do so?*
> : >
> : > *I know that on IE5, windows.selection.createRange ()*
> : > *might help. However, when I finish modifying the selected*
> : > *text, I don't know how to put them back. Using pasteHTML*
> : > *won't work, since there'll be an extra <P> before "TEXT"*
> : > *and an extra after "ITEM". Thus the "Bold TEXT" would*
> : > *be broken into two lines.*
> : >
> : > *Somebody can think of another way else?*

microsoft.public.scripting.jscript: Re: (repost): Very tough question, about HTML manipulation

> : >
> : > *W. Jordan*
> : >
> : >
> : >
> : >
> : >
>
>
>