

Re: REQ – Help in editing a file – Script or Utility

Source:

<http://www.tech-archive.net/Archive/Scripting/microsoft.public.scripting.vbscript/2007-06/msg00148.html>

- *From:* "mayayana" <mayaXXyana1a@xxxxxxxxxxxxxxxxxxxxx>
 - *Date:* Thu, 07 Jun 2007 04:32:30 GMT
-

There are different ways to do that. The biggest issue is that it's a very clunky operation in script. One way would be to just loop through with InStr, concatenating each block of text not between (and). It's more efficient if you can deal with it as an array. Concatenation is costly, and you could get bogged down with the memory usage that requires. As the string gets bigger you're dealing with the need to re-allocate memory for the whole string, plus the new snippet, on each concatenation. I've never been able to find any 1-step method that allows VBS to treat a string as a numeric (or even character) array for processing, but there is a partial solution that should be better than gluing together snippets found via InStr –

The method below seems to be fairly efficient. It uses a "tokenizing" routine instead of a string search. Tokenizing is going through the string 1 character at a time. It's surprisingly fast and allows you to use any level of complexity in a Select Case to edit the string. I tested the script below on a 1.6 MB server log file, with about 7,600 lines, most of which have a section in parentheses. To open the file, remove the parentheses sections, and write the file back to disk, was about 7–8 seconds on my Win98SE Athlon 1660.

Rather than snipping out thousands of small strings for concatenation, it builds a character array, avoiding large memory allocations by just adding characters to the array 1 by 1. At the end it's all put back into a string with Join, which is almost instant.

It would probably be quite a bit more efficient if Split could be used at the beginning to do all the work in arrays, rather than using Mid. But there was an unfortunate oversight when they designed the Split function. It doesn't work like Join. With `s = Join(a, "")` you convert array a to a string very efficiently. But if you do the reverse... `a = Split(s, "")`you just get an array with one element: `a(0) = s !!`

Re: REQ – Help in editing a file – Script or Utility

I wasn't clear about what exactly you want to remove. Your post says everything between ("(and)") but there are no such patterns in the sample. If you want to delete between ("{ and}") you can change the sample script to use {} instead of (). If you want to start the delete at (" you can use two or more booleans to respond only to combinations of characters. Or you can use nested loops. Or you can use nested Ifs, like:

```
Case "("
'-- drop out this next section only if it's ("
If Mid(s1, i2 + 1, 1) = chr(34) then SkipThis = True
```

That's the nice thing about tokenizing routines. There's really no limit to how fine-grained you can make it when every single character gets filtered through a Select Case.

Drop a file onto this script to remove all text from
(to).

```
Dim A1(), s1, s2, i2, LenS, i3, SkipThis, Arg, FSO, TS
```

```
Arg = WScript.Arguments(0)
Set FSO = CreateObject("Scripting.FileSystemObject")
```

```
Set TS = FSO.OpenTextFile(Arg, 1)
s1 = TS.ReadAll
TS.Close
Set TS = Nothing
```

```
LenS = Len(s1)
ReDim A1(LenS)
SkipThis = False
i3 = 0
```

```
For i2 = 1 to LenS
s2 = Mid(s1, i2, 1)
Select Case s2
Case "("
SkipThis = True
Case ")"
SkipThis = False
Case Else
If (SkipThis = False) Then
A1(i3) = s2
i3 = i3 + 1
End If
End Select
Next
```

```
ReDim preserve A1(i3)
s1 = Join(A1, "")
```

```
Set TS = FSO.CreateTextFile(Arg, True)
TS.Write s1
TS.Close
Set TS = Nothing
Set FSO = Nothing
```

```
MsgBox "Done"
```

G'day all

I have to sort through some huge text files and delete most of the content and leave only what is needed. Most of this can be done with "find/replace" from any editor but I'm stuck trying to get rid of two particular components of the file. One is the

line

that is just a hard return

and the other is a string that occurs hundreds of times but is unique in every instance.

An example of the file follows and what I wish to delete is the string between the character "(" and ")".

How to do this with a VBscript or a suggestion of a utility would be great.

Thanks regards

Mark

```
)
("{97AEB36B-9AEA-11D5-BD16-0090272CCB30}")
:ClassName (dynamic_object)
:table (network_objects)
:name (LocalMachine)
)
)
("{EA30D2F2-BC94-4B71-B92F-CC9542CD8C06}")
:ClassName (gateway_ckp)
:object_permissions (
:AdminInfo (
("{7D89199E-2418-4F71-83C5-C0304F9F0897}")
: "#certreq-pki-gen" (false)
: "#pki-host-cert-set" (false)
:ca (ReferenceObject

("{94CEFD0A-46E3-4F63-9007-C15F424F5E48}")
:ClassName (certificate)
```

Re: REQ – Help in editing a file – Script or Utility

```
:Name (internal_ca)
:Table (servers)
("{91788B8E-4A01-420A-87A3-2317DF360EF0}")
)
)
)
:edges ()
:chkpf_uid ("{ECB98557-5E00-442D-B2C1-393DAAA7A780}")
:ClassName (interface)
:netaccess (
:AdminInfo (
("{542CA05E-E2C1-4F0D-B5EB-C0F93314C66A}")
:ClassName (netaccess)
:Table (network_objects)
("{CA6EA914-133D-4D21-B25A-E996D415D8FD}")
)
("{ECB98557-5E00-442D-B2C1-393DAAA7A780}")
:ClassName (interface_security)
)
)
)
:1 (
(
("{90F4F48E-42BA-415A-8281-E4BEBCDF4AA7}")
:ClassName (interface)
```