
Subject: RegExp::ReplaceGlobal() is buggy
Posted by [zsoft](#) on Tue, 11 Jun 2019 22:47:49 GMT
[View Forum Message](#) <> [Reply to Message](#)

I did not investigate it, but multiple replaces are not working well:

```
String res8;

res8 = "ab<a>ZZ</a>";
RegExp("<\\s*a[^>]*>(.*?)<\\s*\\s*a>", RegExp::UNICODE|RegExp::CASELESS)
  .ReplaceGlobal(res8, "");
CHECK_EQUAL("ab", ~res8); //this is OK

res8 = "12<a>ZZ</a>34<a>vv</a>56";
RegExp("<\\s*a[^>]*>(.*?)<\\s*\\s*a>", RegExp::UNICODE|RegExp::CASELESS)
  .ReplaceGlobal(res8, "");
CHECK_EQUAL("123456", ~res8); //buggy
```

The ReplaceGlobal implementation is too hard for me, so I created a very simple but working method for myself to be able to use regular expressions to multiple replaces:

```
String RegExp::ReplaceAll(const String& s, const String& r)
{
  String res;
  int iPosStart;
  int iPosAfterEnd;
  int iFromCopyToRes = 0;
  while(GlobalMatch(s)){
    GetMatchPos(0, iPosStart, iPosAfterEnd);
    //String matched = s.Mid(iPosStart, iPosAfterEnd-iPosStart);//debug
    res << s.Mid(iFromCopyToRes, iPosStart-iFromCopyToRes);
    res << r;
    iFromCopyToRes = iPosAfterEnd;
  }
  if(iFromCopyToRes < s.GetCount()){
    res << s.Mid(iFromCopyToRes, s.GetCount()-iFromCopyToRes);
  }
  return res;
}
```

Test code:

```
String res8;
res8 = "12<a>ZZ</a>34<a>vv</a>56";
res8 = RegExp("<\\s*a[^>]*>(.*?)<\\s*\\s*a>", RegExp::UNICODE|RegExp::CASELESS)
```

```
.ReplaceAll(res8, "");  
CHECK_EQUAL("123456", ~res8);
```

Subject: Re: RegExp::ReplaceGlobal() is buggy
Posted by [mirek](#) on Wed, 12 Jun 2019 08:10:01 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks. Before fixing RegExp, have you considered using
<https://en.cppreference.com/w/cpp/regex> ?

Mirek

Subject: Re: RegExp::ReplaceGlobal() is buggy
Posted by [zsolt](#) on Wed, 12 Jun 2019 11:02:58 GMT
[View Forum Message](#) <> [Reply to Message](#)

Thanks, I didn't know, that it is in the standard now.
Btw, Using RegEx class is much more easy.

Subject: Re: RegExp::ReplaceGlobal() is buggy
Posted by [mirek](#) on Wed, 12 Jun 2019 11:38:18 GMT
[View Forum Message](#) <> [Reply to Message](#)

zsolt wrote on Wed, 12 June 2019 13:02 Thanks, I didn't know, that it is in the standard now.
Btw, Using RegEx class is much more easy.

Well, I hoped that somebody would try and tell me whether to keep maintaining RegExp... "0

Mirek
