Subject: Bug in xml parser?
Posted by aftershock on Fri, 09 Jul 2010 20:55:04 GMT
View Forum Message <> Reply to Message

Hi,

I wonder if there is bug in xml parser because of the following problem.

I use XMLRPC package.
I debuged the package. I can see this.
The server returns an xml which something like this:

<value><string>line1\nline2\nline3\n</string></value>

The parser returns a string which is "line1\nline2\nline3" without the final \n.
My question is .. Is this bug if the xml parser does not return the final \n?

It looks so.

This is how I fixed it:
Add this line to xmlparser class definition:
    void   Preserve(bool b)   { npreserve = b; }

Modified xmlrpc\client.cpp
This is the result of changes around line 65.

```
if(IsNull(response)) {
  faultCode = XMLRPC_CLIENT_HTTP_ERROR;
  faultString = server.GetError();
  error = "Http request failed: " + faultString;
  LLOG(error);
  return ErrorValue(error);
 }
 XmlParser p(response);
 p.Preserve(true); //<-- new line
 try {
```

Subject: Re: Bug in xml parser?
Posted by mirek on Sat, 10 Jul 2010 10:37:02 GMT
View Forum Message <> Reply to Message

aftershock wrote on Fri, 09 July 2010 16:55Hi,

I wonder if there is bug in xml parser because of the following problem.

I use XMLRPC package.
I debuged the package. I can see this.
The server returns an xml which something like this:

<value><string>line1\nline2\nline3\n</string></value>

The parser returns a string which is "line1\nline2\nline3" without the final \n.
My question is .. Is this bug if the xml parser does not return the final \n?

Well, the problem is somewhat fuzzy...

Quote:
This is how I fixed it:
Add this line to xmlparser class definition:
    void   Preserve(bool b)   { npreserve = b; }

I have implemented very similar fix, only instead of using npreserve I have used another bool attribute.

Mirek

Subject: Re: Bug in xml parser?
Posted by aftershock on Sun, 11 Jul 2010 06:11:04 GMT
View Forum Message <> Reply to Message

What boolean attribute?

Subject: Re: Bug in xml parser?
Posted by mirek on Sun, 11 Jul 2010 08:01:22 GMT
View Forum Message <> Reply to Message

aftershock wrote on Sun, 11 July 2010 02:11What boolean attribute?

"attribute" is alternate name for "instance member variable"...

What about the proxy?

Mirek

Subject: Re: Bug in xml parser?
Posted by mirek on Fri, 23 Jul 2010 08:46:33 GMT
View Forum Message <> Reply to Message

aftershock wrote on Sun, 18 July 2010 05:12luzr wrote on Sun, 18 July 2010 10:03aftershock wrote on Mon, 12 July 2010 12:27I think this is the fix

```
void XmlParser::Next()
{
 nattr.Clear();
 nattr1 = nattrval1 = Null;
 if(empty_tag) {
  empty_tag = false;
  type = XML_END;
  return;
 }
 text.Clear();
 type = XML_EOF;
 if (!(npreserve || preserveall)) //<-- this should be added
   SkipWhites();
```

The problem is the initial whitespace and EOL is thrown away which is not good.
Your other fix seems to be ok.

Have a good day!

Marton

Hi,

I had to revert this patch as it completely breaks any existing XML code.

I suggest you to send me some testcase demonstrating XML that needs XmlParser to be fixed, I will try to find some workaround.

Mirek
It was not my intention Really? How does it break them?

E.g
<value> myvalue</value>
This should return " myvalue" if it is read..

Marton

That is not a problem. The trouble is that when active, it will not skip any whitespaces in between tags, so:

<tag1>
  <tag2>

cannot be parsed with

  if(Tag("tag1") && Tag("tag2"))

as there is now Text element in between....

Sure, perhaps we might try to alter parser philosophy, but at this point I think that would be dangerous and we should try to find another solution first...

Mirek

(reposting back to forum to eventually gather more comments).

Subject: Re: Bug in xml parser?
Posted by aftershock on Sat, 11 Sep 2010 09:50:23 GMT
View Forum Message <> Reply to Message

Just one comment...
Parsing whitespace as text between tags is correct formally,right?

Subject: Re: Bug in xml parser?
Posted by mirek on Sat, 25 Sep 2010 13:25:14 GMT
View Forum Message <> Reply to Message

aftershock wrote on Sat, 11 September 2010 05:50Just one comment...
Parsing whitespace as text between tags is correct formally,right?

As I understand it, yes. But pain in the ass for any "normal" XML...
