Redmine - Defect #1591

http links containing parentheses fail to reder correctly

2008-07-07 05:07 - ChunChang (Nagaharu) Lo

Status:	Closed		Start date:	2008-07-07
Priority:	Normal		**Due date:**	
Assignee:	Jean-Philippe Lang		**% Done:**	0%
Category:			**Estimated time:**	0.00 hour
Target version:	0.8			
Resolution:	Fixed		**Affected version:**	0.7.3

Description
for example: http://en.wikipedia.org/wiki/Reflection_(computer_science) http://en.wikipedia.org/wiki/Introspection_(computer_science) http://en.wikipedia.org/wiki/Introspector_(program) will be rendered without right parenthesis.

Related issues:			
Has duplicate Redmine - Defect #1831: Textile bug in forming URL with last ')'...		**Closed**	**2008-08-29**

Associated revisions

Revision 1871 - 2008-09-17 18:48 - Jean-Philippe Lang

Fixed: http links containing parentheses fail to reder correctly (#1591). Patch by Paul Rivier.

Revision 1874 - 2008-09-17 20:42 - Nicolas Chuche

r18596@gaspard (orig r1860): nbc | 2008-09-14 21:03:46 +0200
bugfix
r18597@gaspard (orig r1861): winterheart | 2008-09-15 17:14:34 +0200
#1902, translation for zh-tw
r18598@gaspard (orig r1862): winterheart | 2008-09-15 17:16:53 +0200
#1907, translation for zh
r18599@gaspard (orig r1863): winterheart | 2008-09-15 17:19:51 +0200
fixed #1905, patch for Hungarian language
r18600@gaspard (orig r1864): winterheart | 2008-09-15 17:22:53 +0200
Minor typo, fixed #1897, thank Denis Tomashenko for reporting.
r18601@gaspard (orig r1865): winterheart | 2008-09-15 18:07:30 +0200
Catalan translation (#1822), thanks to Joan Duran for contribuition. Some strings has wrong quoting, I fixed that.
r18602@gaspard (orig r1866): nbc | 2008-09-15 21:37:43 +0200 * reposman can create git repository with "--scm git" option * light refactoring
r18603@gaspard (orig r1867): jplang | 2008-09-16 23:54:53 +0200
Use RDoc.usage
r18604@gaspard (orig r1868): jplang | 2008-09-16 23:56:02 +0200
mailhandler: fixes exit status and adds an explicit message if response code is 403.
r18605@gaspard (orig r1869): winterheart | 2008-09-17 17:31:35 +0200
Patch #1909, updates for ru.yml
r18606@gaspard (orig r1870): jplang | 2008-09-17 18:39:23 +0200
Render the commit changes list as a tree (#1896).
r18607@gaspard (orig r1871): jplang | 2008-09-17 18:48:04 +0200
Fixed: http links containing parentheses fail to reder correctly (#1591). Patch by Paul Rivier.
r18608@gaspard (orig r1872): jplang | 2008-09-17 19:18:05 +0200
Removes unused image references in stylesheets (#1914).
r18609@gaspard (orig r1873): jplang | 2008-09-17 19:23:08 +0200
Fixed custom query sidebar links broken by r1797 (#1899).

History

#1 - 2008-07-10 18:31 - Pierre Paysant-Le Roux

- File patch_links_with_parenthesis.patch added

Here is a patch that should solve the problem. The regexp come from Rails with some modifications.

#2 - 2008-07-13 12:42 - Jean-Philippe Lang

Thanks for the patch.
But it would be nice to fix RedCloth#inline_textile_link as well.

#3 - 2008-08-25 16:03 - Jean-Philippe Lang

This patch doesn't work well with links inside parenthesis. Eg. (http://www.foo.bar/Test_foobar).
Trailing parenthesis is included in the url.

#4 - 2008-08-25 17:53 - Paul Rivier

- Assignee set to Paul Rivier

Jean-Philippe Lang wrote:

> This patch doesn't work well with links inside parenthesis. Eg. (http://www.foo.bar/Test_foobar).
> Trailing parenthesis is included in the url.

This is not an easy problem. Should we consider that parenthesis inside url should alway be balanced ? I don't think we should.
Then should we treat the particular case of a url with a preceding (and a trailing) ?

#5 - 2008-08-25 18:01 - Paul Rivier

Paul Rivier wrote:

> This is not an easy problem. Should we consider that parenthesis inside url should alway be balanced ? I don't think we should.
> Then should we treat the particular case of a url with a preceding (and a trailing) ?

here is a naive patch for the later.

```
   def inline_auto_link(text)
       text.gsub!(AUTO_LINK_RE) do
         all, a, b, c, d = $&, $1, $2, $3, $4
         if a =~ /<a\s/i || a =~ /![<>=]?/
         else
################### PATCH BELOW
           if ( a[-1] == ?( and c[-1]==?) ) # looks like url is put inside parenthesis
             c=c[0..-2]
             d=")"+d
           end
################# PATCH ABOVE
           text = b + c
           %(#{a}<a class="external" href="#{b=="www."?"http://www.":b}#{c}">#{text}</a>#{d})
         end
       end
     end
```

#6 - 2008-08-25 18:04 - Paul Rivier

Jean-Philippe Lang wrote:

> Thanks for the patch.
> But it would be nice to fix RedCloth#inline_textile_link as well.

What's the policy for this kind of "bundled-librairy" related problem ? Should we patch against bundled files ? Should we try to monkey-patch from outside ?

#7 - 2008-08-25 21:04 - Jean-Philippe Lang

> Should we consider that parenthesis inside url should alway be balanced ? I don't think we should.

I don't know which is the best solution but I'd really like that something like this works: (see: http://example.net).
I think we see this more often than a link with a single closing parenthesis at the end.
WDYT ?

#8 - 2008-08-25 21:06 - Jean-Philippe Lang

> Should we patch against bundled files ?

Yes, you can.

redcloth.rb is already patched: http://www.redmine.org/repositories/changes/redmine/trunk/lib/redcloth.rb

#9 - 2008-08-25 21:27 - Paul Rivier

Jean-Philippe Lang wrote:

> I don't know which is the best solution but I'd really like that something like this works: (see: http://example.net).
> I think we see this more often than a link with a single closing parenthesis at the end.
> WDYT ?

For sure I'd like this as well, but we first have to formulate clearly what we want. For the moment, it's like :

```
"If the url ends a sentence in parenthesis, don't include the closing parenthesis in the URL"
```

That would involve matching the whole sentence, before url to see if it starts with '(' and after to see if there isn't an other closing ')' later, like in "(see: http://fr.example.net/wiki/clavier_(musique) to read more)". I think we will have hard time doing so :/

Maybe we can do that :

```
"If the url last char is closing ')' that is unbalanced in the url itself, take the ')' as external to the u
rl
```

This seems doable, does not require to scan far around the url itself, and should handle :

- (see: http://example.net)
- (see: http://fr.example.net/wiki/clavier_(musique) to read more)
- (see: http://fr.example.net/wiki/clavier_(musique))

WDYT ?

#10 - 2008-08-25 21:44 - Jean-Philippe Lang

> "If the url last char is closing ')' that is unbalanced in the url itself, take the ')' as external to the url

This is fine for me.

#11 - 2008-08-26 12:01 - Paul Rivier

- File urls_fix.patch added

Jean-Philippe Lang wrote:

> > "If the url last char is closing ')' that is unbalanced in the url itself, take the ')' as external to the url
>
> This is fine for me.

so please find attached patch against current trunk. It's fairly minor, and does not change the (already fairly complex) regexp for url except concerning the parenthesis. Tests included.

#12 - 2008-09-04 11:55 - Paul Rivier

jean philippe, are you reasonably happy with this patch ? If not, could you explain how you think it should be improved please.

#13 - 2008-09-17 15:45 - Paul Rivier

- Status changed from New to 7

- Assignee changed from Paul Rivier to Jean-Philippe Lang

What's up with this patch please ?
Jean Philippe, you are the new assignee, as you may want to apply it or not, or close the issue.

#14 - 2008-09-17 18:43 - Jean-Philippe Lang

- Status changed from 7 to Closed

- Target version set to 0.8

- Affected version (unused) set to 0.7.3

- Resolution set to Fixed

- Affected version set to 0.7.3

Paul, I didn't see your new patch, sorry.
Anyway, everything looks fine and it's now committed in [r1871](#).

Thanks !

Files

patch_links_with_parenthesis.patch	5.01 KB	2008-07-10	Pierre Paysant-Le Roux
urls_fix.patch	4.37 KB	2008-08-26	Paul Rivier