

Re: match nested tags

Source: <http://coding.derkeiler.com/Archive/Perl/comp.lang.perl.modules/2006-05/msg00020.html>

- *From:* "DJ Stunks" <DJStunks@xxxxxxxxx>
 - *Date:* 3 May 2006 14:20:19 -0700
-

Hey guys, I'm not getting any responses over at perl.beginners so I thought I'd cross post this here to see if anyone has any ideas.

Here's the original message:

FangQ wrote:

```
> hi
>
> is there a simple way using regular expression to find nested tags?
>
> for example, the string is:
>
> {{ {A} this is part A of the document
> {{ {A.1} this is part A1 }}
> }}
>
> I want to define a function findtag("A") to give me
>
> this is part A of the document
> {{ {A.1} this is part A1 }}
>
>
> and findtag("A.1") to give me
>
> this is part A1
>
> can anyone give some hint?
> thanks
```

I thought this sounded like a prime candidate for Parse::RecDescent, but I can't get the nested nature of the part(s) to work.

Here's my first crack at it, but it doesn't parse. I monkeyed with it for a while, but to no avail.

I did note, however, that in the Parse::RecDescent FAQ, Pastor Conway suggests using Text::Balanced to extract nested parenthesis. I tried that too, but again, no luck.

Re: match nested tags

I'd be interested to see if anyone here has a suggestion for this problem. Thanks in advance.

```
#!/usr/bin/perl

use strict;
use warnings;

use Data::Dumper;
use Parse::RecDescent;

my $grammar = <<'EO_GRAMMAR';
<autotree>

document : '{' part(s) '}'

part : part_id part_text part(s?)
part_id : '[' / [^] + '/'
part_text : /.+ /s

EO_GRAMMAR

my $parser = Parse::RecDescent->new($grammar)
or die "Could not parse grammar: $@";

my $document = do {local $/; <DATA>};

my $doc_ref = $parser->document($document)
or die "Invalid document";

print Dumper $doc_ref;

__DATA__
{{ {A} this is part A of the document
{{ {A.1} this is part A1 }}
}}

__END__

-jp
```