**configs详解——之selector**

selector是页面元素选择器类，下面介绍此类可以调用的方法

**select($html, $selector, $selector\_type = 'xpath')**

@param $html 需筛选的网页内容  
@param $selector 选择器规则  
@param $selector\_type 选择器类型: xpath、regex、css, 默认为xpath选择类型

栗子1:

通过xpath选择器提取网页内容的标题

$html = requests::get("http://www.epooll.com/archives/806/");

$data = selector::select($html, "//div[contains(@class,'page-header')]//h1//a");

var\_dump($data);

栗子2:

通过css选择器提取网页内容的标题

$html = requests::get("http://www.epooll.com/archives/806/");

$data = selector::select($html, ".page-header > h1 > a", "css");

var\_dump($data);

栗子3:

通过正则匹配提取网页内容的标题

$html = requests::get("http://www.epooll.com/archives/806/");

$data = selector::select($html, '@<title>(.\*?)</title>@', "regex");

var\_dump($data);

**remove($html, $selector, $selector\_type = 'xpath')**

@param $html 需过滤的网页内容  
@param $selector 选择器规则  
@param $selector\_type 选择器类型: xpath、regex、css, 默认为xpath选择类型

举个例子:

$html =<<<STR

<div id="demo">

aaa

<span class="tt">bbb</span>

<span>ccc</span>

<p>ddd</p>

</div>

STR;

// 获取id为demo的div内容

$html = selector::select($html, "//div[contains(@id,'demo')]");

// 在上面获取内容基础上，删除class为tt的span标签

$data = selector::remove($html, "//span[contains(@class,'tt')]");

print\_r($data);