PHP Simple HTML DOM Parser

Index

API Reference

Helper functions

Name	Description
object str_get_html (string $content)	Creates a DOM object from a string.
object file_get_html (string $filename)	Creates a DOM object from a file or a URL.

DOM methods & properties

Name	Description
void __construct ([string $filename])	Constructor, set the filename parameter will automatically load the contents, either text or file/url.
string plaintext	Returns the contents extracted from HTML.
void clear ()	Clean up memory.
void load (string $content)	Load contents from a **string**.
string save ([string $filename])	Dumps the internal DOM tree back into a **string**. If the $filename is set, result string will save to file.
void load_file (string $filename)	Load contents from a from a file or a URL.
void set_callback (string $function_name)	Set a callback function.
mixed find (string $selector [, int $index])	Find elements by the CSS selector. Returns the Nth element **object** if **index** is set, otherwise return an **array** of object.

Element methods & properties

Name	Description
string [attribute]	Read or write element's attribure value.
string tag	Read or write the tag name of element.
string outertext	Read or write the outer HTML text of element.
string innertext	Read or write the inner HTML text of element.
string plaintext	Read or write the plain text of element.
mixed find (string $selector [, int $index])	Find children by the CSS selector. Returns the Nth element **object** if **index** is set, otherwise, return an **array** of object.

DOM traversing

Name	Description
mixed $e->children ([int $index])	Returns the Nth child **object** if **index** is set, otherwise return an **array** of children.
element $e->parent ()	Returns the parent of element.
element $e->first_child ()	Returns the first child of element, or **null** if not found.
element $e->last_child ()	Returns the last child of element, or **null** if not found.
element $e->next_sibling ()	Returns the next sibling of element, or **null** if not found.

element $e->prev_sibling ()	Returns the previous sibling of element, or **null** if not found.

Camel naming convertions

You can also call methods with W3C STANDARD camel naming convertions.

Method	Mapping
array $e->getAllAttributes ()	array $e->attr
string $e->getAttribute ($name)	string $e->attribute
void $e->setAttribute ($name, $value)	void $value = $e->attribute
bool $e->hasAttribute ($name)	bool isset($e->attribute)
void $e->removeAttribute ($name)	void $e->attribute = null
element $e->getElementById ($id)	mixed $e->find ("#$id", 0)
mixed $e->getElementsById ($id [,$index])	mixed $e->find ("#$id" [, int $index])
element $e->getElementByTagName ($name)	mixed $e->find ($name, 0)
mixed $e->getElementsByTagName ($name [, $index])	mixed $e->find ($name [, int $index])
element $e->parentNode ()	element $e->parent ()
mixed $e->childNodes ([$index])	mixed $e->children ([int $index])
element $e->firstChild ()	element $e->first_child ()
element $e->lastChild ()	element $e->last_child ()
element $e->nextSibling ()	element $e->next_sibling ()
element $e->previousSibling ()	element $e->prev_sibling ()

Author: S.C. Chen (me578022@gmail.com)

Original idea is from Jose Solorzano's HTML Parser for PHP 4.

Contributions by: Contributions by: Yousuke Kumakura, Vadim Voituk, Antcs