| | [**Overview**](http://docs.google.com/overview-summary.html) | [**Package**](http://docs.google.com/package-summary.html) | **Class** | [**Use**](http://docs.google.com/class-use/DocumentParser.html) | [**Tree**](http://docs.google.com/package-tree.html) | [**Deprecated**](http://docs.google.com/deprecated-list.html) | [**Index**](http://docs.google.com/index-files/index-1.html) | [**Help**](http://docs.google.com/help-doc.html) | | --- | --- | --- | --- | --- | --- | --- | --- | | | ***Java™ Platform***  ***Standard Ed. 6*** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [**PREV CLASS**](http://docs.google.com/javax/swing/text/html/parser/ContentModel.html)   [**NEXT CLASS**](http://docs.google.com/javax/swing/text/html/parser/DTD.html) | [**FRAMES**](http://docs.google.com/index.html?javax/swing/text/html/parser/DocumentParser.html)    [**NO FRAMES**](http://docs.google.com/DocumentParser.html)     [**All Classes**](http://docs.google.com/allclasses-noframe.html) |
| SUMMARY: NESTED | [FIELD](#2et92p0) | [CONSTR](#3dy6vkm) | [METHOD](#1t3h5sf) | DETAIL: FIELD | [CONSTR](#17dp8vu) | [METHOD](#26in1rg) |

## **javax.swing.text.html.parser**

Class DocumentParser

[java.lang.Object](http://docs.google.com/java/lang/Object.html)  
 [javax.swing.text.html.parser.Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)  
 **javax.swing.text.html.parser.DocumentParser**

**All Implemented Interfaces:** [DTDConstants](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html)

public class **DocumentParser**extends [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

A Parser for HTML Documents (actually, you can specify a DTD, but you should really only use this class with the html dtd in swing). Reads an InputStream of HTML and invokes the appropriate methods in the ParserCallback class. This is the default parser used by HTMLEditorKit to parse HTML url's.

This will message the callback for all valid tags, as well as tags that are implied but not explicitly specified. For example, the html string (<p>blah) only has a p tag defined. The callback will see the following methods:

1. *handleStartTag(html, ...)*
2. *handleStartTag(head, ...)*
3. *handleEndTag(head)*
4. *handleStartTag(body, ...)*
5. handleStartTag(p, ...)
6. handleText(...)
7. *handleEndTag(p)*
8. *handleEndTag(body)*
9. *handleEndTag(html)*

The items in *italic* are implied, that is, although they were not explicitly specified, to be correct html they should have been present (head isn't necessary, but it is still generated). For tags that are implied, the AttributeSet argument will have a value of Boolean.TRUE for the key HTMLEditorKit.ParserCallback.IMPLIED.

HTML.Attributes defines a type safe enumeration of html attributes. If an attribute key of a tag is defined in HTML.Attribute, the HTML.Attribute will be used as the key, otherwise a String will be used. For example <p foo=bar class=neat> has two attributes. foo is not defined in HTML.Attribute, where as class is, therefore the AttributeSet will have two values in it, HTML.Attribute.CLASS with a String value of 'neat' and the String key 'foo' with a String value of 'bar'.

The position argument will indicate the start of the tag, comment or text. Similiar to arrays, the first character in the stream has a position of 0. For tags that are implied the position will indicate the location of the next encountered tag. In the first example, the implied start body and html tags will have the same position as the p tag, and the implied end p, html and body tags will all have the same position.

As html skips whitespace the position for text will be the position of the first valid character, eg in the string '\n\n\nblah' the text 'blah' will have a position of 3, the newlines are skipped.

For attributes that do not have a value, eg in the html string <foo blah> the attribute blah does not have a value, there are two possible values that will be placed in the AttributeSet's value:

* If the DTD does not contain an definition for the element, or the definition does not have an explicit value then the value in the AttributeSet will be HTML.NULL\_ATTRIBUTE\_VALUE.
* If the DTD contains an explicit value, as in: <!ATTLIST OPTION selected (selected) #IMPLIED> this value from the dtd (in this case selected) will be used.

Once the stream has been parsed, the callback is notified of the most likely end of line string. The end of line string will be one of \n, \r or \r\n, which ever is encountered the most in parsing the stream.

| **Field Summary** | |
| --- | --- |

| **Fields inherited from class javax.swing.text.html.parser.**[**Parser**](http://docs.google.com/javax/swing/text/html/parser/Parser.html) |
| --- |
| [dtd](http://docs.google.com/javax/swing/text/html/parser/Parser.html#dtd), [strict](http://docs.google.com/javax/swing/text/html/parser/Parser.html#strict) |

| **Fields inherited from interface javax.swing.text.html.parser.**[**DTDConstants**](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html) |
| --- |
| [ANY](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#ANY), [CDATA](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#CDATA), [CONREF](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#CONREF), [CURRENT](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#CURRENT), [DEFAULT](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#DEFAULT), [EMPTY](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#EMPTY), [ENDTAG](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#ENDTAG), [ENTITIES](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#ENTITIES), [ENTITY](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#ENTITY), [FIXED](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#FIXED), [GENERAL](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#GENERAL), [ID](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#ID), [IDREF](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#IDREF), [IDREFS](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#IDREFS), [IMPLIED](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#IMPLIED), [MD](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#MD), [MODEL](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#MODEL), [MS](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#MS), [NAME](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NAME), [NAMES](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NAMES), [NMTOKEN](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NMTOKEN), [NMTOKENS](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NMTOKENS), [NOTATION](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NOTATION), [NUMBER](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NUMBER), [NUMBERS](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NUMBERS), [NUTOKEN](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NUTOKEN), [NUTOKENS](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#NUTOKENS), [PARAMETER](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#PARAMETER), [PI](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#PI), [PUBLIC](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#PUBLIC), [RCDATA](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#RCDATA), [REQUIRED](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#REQUIRED), [SDATA](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#SDATA), [STARTTAG](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#STARTTAG), [SYSTEM](http://docs.google.com/javax/swing/text/html/parser/DTDConstants.html#SYSTEM) |

| **Constructor Summary** | |
| --- | --- |
| [**DocumentParser**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#DocumentParser(javax.swing.text.html.parser.DTD))([DTD](http://docs.google.com/javax/swing/text/html/parser/DTD.html) dtd) |

| **Method Summary** | |
| --- | --- |
| protected  void | [**handleComment**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleComment(char%5B%5D))(char[] text)            Called when an HTML comment is encountered. |
| protected  void | [**handleEmptyTag**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleEmptyTag(javax.swing.text.html.parser.TagElement))([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)            Handle Empty Tag. |
| protected  void | [**handleEndTag**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleEndTag(javax.swing.text.html.parser.TagElement))([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)            Handle End Tag. |
| protected  void | [**handleError**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleError(int,%20java.lang.String))(int ln, [String](http://docs.google.com/java/lang/String.html) errorMsg)            An error has occurred. |
| protected  void | [**handleStartTag**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleStartTag(javax.swing.text.html.parser.TagElement))([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)            Handle Start Tag. |
| protected  void | [**handleText**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#handleText(char%5B%5D))(char[] data)            Handle Text. |
| void | [**parse**](http://docs.google.com/javax/swing/text/html/parser/DocumentParser.html#parse(java.io.Reader,%20javax.swing.text.html.HTMLEditorKit.ParserCallback,%20boolean))([Reader](http://docs.google.com/java/io/Reader.html) in, [HTMLEditorKit.ParserCallback](http://docs.google.com/javax/swing/text/html/HTMLEditorKit.ParserCallback.html) callback, boolean ignoreCharSet) |

| **Methods inherited from class javax.swing.text.html.parser.**[**Parser**](http://docs.google.com/javax/swing/text/html/parser/Parser.html) |
| --- |
| [endTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#endTag(boolean)), [error](http://docs.google.com/javax/swing/text/html/parser/Parser.html#error(java.lang.String)), [error](http://docs.google.com/javax/swing/text/html/parser/Parser.html#error(java.lang.String,%20java.lang.String)), [error](http://docs.google.com/javax/swing/text/html/parser/Parser.html#error(java.lang.String,%20java.lang.String,%20java.lang.String)), [error](http://docs.google.com/javax/swing/text/html/parser/Parser.html#error(java.lang.String,%20java.lang.String,%20java.lang.String,%20java.lang.String)), [flushAttributes](http://docs.google.com/javax/swing/text/html/parser/Parser.html#flushAttributes()), [getAttributes](http://docs.google.com/javax/swing/text/html/parser/Parser.html#getAttributes()), [getCurrentLine](http://docs.google.com/javax/swing/text/html/parser/Parser.html#getCurrentLine()), [getCurrentPos](http://docs.google.com/javax/swing/text/html/parser/Parser.html#getCurrentPos()), [handleEOFInComment](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleEOFInComment()), [handleTitle](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleTitle(char%5B%5D)), [makeTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#makeTag(javax.swing.text.html.parser.Element)), [makeTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#makeTag(javax.swing.text.html.parser.Element,%20boolean)), [markFirstTime](http://docs.google.com/javax/swing/text/html/parser/Parser.html#markFirstTime(javax.swing.text.html.parser.Element)), [parse](http://docs.google.com/javax/swing/text/html/parser/Parser.html#parse(java.io.Reader)), [parseDTDMarkup](http://docs.google.com/javax/swing/text/html/parser/Parser.html#parseDTDMarkup()), [parseMarkupDeclarations](http://docs.google.com/javax/swing/text/html/parser/Parser.html#parseMarkupDeclarations(java.lang.StringBuffer)), [startTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#startTag(javax.swing.text.html.parser.TagElement)) |

| **Methods inherited from class java.lang.**[**Object**](http://docs.google.com/java/lang/Object.html) |
| --- |
| [clone](http://docs.google.com/java/lang/Object.html#clone()), [equals](http://docs.google.com/java/lang/Object.html#equals(java.lang.Object)), [finalize](http://docs.google.com/java/lang/Object.html#finalize()), [getClass](http://docs.google.com/java/lang/Object.html#getClass()), [hashCode](http://docs.google.com/java/lang/Object.html#hashCode()), [notify](http://docs.google.com/java/lang/Object.html#notify()), [notifyAll](http://docs.google.com/java/lang/Object.html#notifyAll()), [toString](http://docs.google.com/java/lang/Object.html#toString()), [wait](http://docs.google.com/java/lang/Object.html#wait()), [wait](http://docs.google.com/java/lang/Object.html#wait(long)), [wait](http://docs.google.com/java/lang/Object.html#wait(long,%20int)) |

| **Constructor Detail** |
| --- |

### DocumentParser

public **DocumentParser**([DTD](http://docs.google.com/javax/swing/text/html/parser/DTD.html) dtd)

| **Method Detail** |
| --- |

### parse

public void **parse**([Reader](http://docs.google.com/java/io/Reader.html) in,  
 [HTMLEditorKit.ParserCallback](http://docs.google.com/javax/swing/text/html/HTMLEditorKit.ParserCallback.html) callback,  
 boolean ignoreCharSet)  
 throws [IOException](http://docs.google.com/java/io/IOException.html)

**Throws:** [IOException](http://docs.google.com/java/io/IOException.html)

### handleStartTag

protected void **handleStartTag**([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)

Handle Start Tag.

**Overrides:**[handleStartTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleStartTag(javax.swing.text.html.parser.TagElement)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

### handleComment

protected void **handleComment**(char[] text)

**Description copied from class:** [**Parser**](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleComment(char%5B%5D)) Called when an HTML comment is encountered.

**Overrides:**[handleComment](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleComment(char%5B%5D)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

### handleEmptyTag

protected void **handleEmptyTag**([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)  
 throws [ChangedCharSetException](http://docs.google.com/javax/swing/text/ChangedCharSetException.html)

Handle Empty Tag.

**Overrides:**[handleEmptyTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleEmptyTag(javax.swing.text.html.parser.TagElement)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html) **Throws:** [ChangedCharSetException](http://docs.google.com/javax/swing/text/ChangedCharSetException.html)

### handleEndTag

protected void **handleEndTag**([TagElement](http://docs.google.com/javax/swing/text/html/parser/TagElement.html) tag)

Handle End Tag.

**Overrides:**[handleEndTag](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleEndTag(javax.swing.text.html.parser.TagElement)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

### handleText

protected void **handleText**(char[] data)

Handle Text.

**Overrides:**[handleText](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleText(char%5B%5D)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

### handleError

protected void **handleError**(int ln,  
 [String](http://docs.google.com/java/lang/String.html) errorMsg)

**Description copied from class:** [**Parser**](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleError(int,%20java.lang.String)) An error has occurred.

**Overrides:**[handleError](http://docs.google.com/javax/swing/text/html/parser/Parser.html#handleError(int,%20java.lang.String)) in class [Parser](http://docs.google.com/javax/swing/text/html/parser/Parser.html)

| | [**Overview**](http://docs.google.com/overview-summary.html) | [**Package**](http://docs.google.com/package-summary.html) | **Class** | [**Use**](http://docs.google.com/class-use/DocumentParser.html) | [**Tree**](http://docs.google.com/package-tree.html) | [**Deprecated**](http://docs.google.com/deprecated-list.html) | [**Index**](http://docs.google.com/index-files/index-1.html) | [**Help**](http://docs.google.com/help-doc.html) | | --- | --- | --- | --- | --- | --- | --- | --- | | | ***Java™ Platform***  ***Standard Ed. 6*** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| [**PREV CLASS**](http://docs.google.com/javax/swing/text/html/parser/ContentModel.html)   [**NEXT CLASS**](http://docs.google.com/javax/swing/text/html/parser/DTD.html) | [**FRAMES**](http://docs.google.com/index.html?javax/swing/text/html/parser/DocumentParser.html)    [**NO FRAMES**](http://docs.google.com/DocumentParser.html)     [**All Classes**](http://docs.google.com/allclasses-noframe.html) |
| SUMMARY: NESTED | [FIELD](#2et92p0) | [CONSTR](#3dy6vkm) | [METHOD](#1t3h5sf) | DETAIL: FIELD | [CONSTR](#17dp8vu) | [METHOD](#26in1rg) |

[Submit a bug or feature](http://bugs.sun.com/services/bugreport/index.jsp)

For further API reference and developer documentation, see [Java SE Developer Documentation](http://docs.google.com/webnotes/devdocs-vs-specs.html). That documentation contains more detailed, developer-targeted descriptions, with conceptual overviews, definitions of terms, workarounds, and working code examples.

Copyright 2006 Sun Microsystems, Inc. All rights reserved. Use is subject to [license terms](http://docs.google.com/legal/license.html). Also see the [documentation redistribution policy](http://java.sun.com/docs/redist.html).