URLEncoder与URLDecoder

**URLDecoder** Utility class for HTML form decoding.

**URLEncoder** Utility class for HTML form encoding.

这是用于HTML表单编码与解码的工具类。

# URLEncoder

## 继承关系

public class **URLEncoder** extends Object



## 功能介绍

Utility class for HTML form encoding. This class contains **static methods** for converting a String to the **application/x-www-form-urlencoded MIME format**. For more information about HTML form encoding, consult the HTML specification.

When encoding a String, the following rules apply:

* The alphanumeric characters "a" through "z", "A" through "Z" and "0" through "9" remain the same.
* The special characters ".", "-", "\*", and "\_" remain the same.
* The space character " " is converted into a plus sign "+".
* All other characters are unsafe and are first converted into one or more bytes using some encoding scheme. Then each byte is represented **by the 3-character string "%xy",** where xy is the two-digit hexadecimal representation of the byte. **The recommended encoding scheme to use is UTF-8.** However, for compatibility reasons, if an encoding is not specified, then the default encoding of the platform is used.

For example using UTF-8 as the encoding scheme the string "The string ü@foo-bar" would get converted to "The+string+%C3%BC%40foo-bar" because in UTF-8 the character ü is encoded as two bytes C3 (hex) and BC (hex), and the character @ is encoded as one byte 40 (hex).

Since: JDK1.0

## 唯一的静态方法：encode(s,enc)

public static String **encode**(String s,String enc) throws UnsupportedEncodingException

Translates a string into application/x-www-form-urlencoded format using a specific encoding scheme. This method uses the supplied encoding scheme to obtain the bytes for unsafe characters.

Note: The World Wide Web Consortium Recommendation states that UTF-8 should be used. Not doing so may introduce incompatibilities.

Parameters:

s - String to be translated.

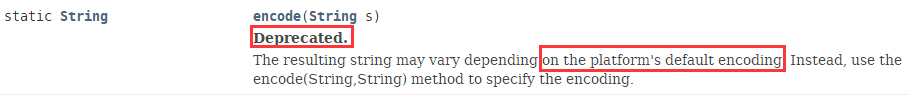
enc - The name of a supported character encoding.

Returns: the translated String.

Throws: UnsupportedEncodingException - If the named encoding is not supported

Since: 1.4

See Also: **URLDecoder.decode(java.lang.String, java.lang.String)**



# URLDecoder

## 继承关系

public class **URLDecoder** extends **Object**



## 功能特性

Utility class for HTML form decoding. This class contains static methods for decoding a String from the application/x-www-form-urlencoded MIME format.

The conversion process is the reverse of that used by the **URLEncoder** class. It is assumed that all characters in the encoded string are one of the following: "a" through "z", "A" through "Z", "0" through "9", and "-", "\_", ".", and "\*". The character "%" is allowed but is interpreted as the start of *a special escaped sequence*.

The following rules are applied in the conversion:

* The alphanumeric characters "a" through "z", "A" through "Z" and "0" through "9" remain the same.
* The special characters ".", "-", "\*", and "\_" remain the same.
* The plus sign "+" is converted into a space character " " .
* A sequence of the form "%xy" will be treated as representing a byte where xy is the two-digit hexadecimal representation of the 8 bits. Then, all substrings that contain one or more of these byte sequences consecutively will be replaced by the character(s) whose encoding would result in those consecutive bytes. The encoding scheme used to decode these characters may be specified, or if unspecified, the default encoding of the platform will be used.

There are two possible ways in which this decoder could deal **with illegal strings**. It could either leave illegal characters alone or it could throw an IllegalArgumentException. Which approach the decoder takes is left to the implementation.

## 唯一的静态方法decode(s,enc)

public static String **decode**(String s, String enc) throws UnsupportedEncodingException

Decodes a application/x-www-form-urlencoded string using a specific encoding scheme. The supplied encoding is used to determine what characters are represented by any consecutive sequences of the form "%xy".

Note: The World Wide Web Consortium Recommendation states that **UTF-8** should be used. Not doing so may introduce incompatibilities.

Parameters: s - the String to decode

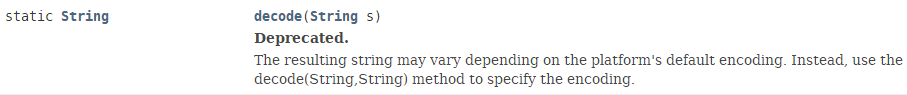
enc - The name of a supported character encoding.

Returns: the newly decoded String

Throws: UnsupportedEncodingException - If character encoding needs to be consulted, but named character encoding is not supported

Since: 1.4

See Also: URLEncoder.encode(java.lang.String, java.lang.String)



# URL编码与解码示例

String s = "我爱我的国家!China";

String urlEncodes = URLEncoder.encode(s,"UTF-8");

System.out.println("urlEncodes = " + urlEncodes);

String urlDecodes = URLDecoder.decode(urlEncodes,"UTF-8");

System.out.println("urlDecodes = " + urlDecodes);

结果：

