**控制字符列表**

来源：http://ascii-table.com/control-chars.php

**Control characters**

**Control codes, C0 controls**

[**Control Characters**](http://ascii-table.com/control-chars.php)  |  [1963 Historical Table](http://ascii-table.com/control-chars-1963.php)

The following document lists the **control characters** in **[Ascii](http://ascii-table.com/ascii.php)** and in newer character code standards like [**Unicode™**](http://ascii-table.com/unicode.php), which is compatible with Ascii in the Ascii code range (positions 0 through 127).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Keys** | **Dec** | **Hex** | **Abbr.** | **Name** | [**Type**](http://ascii-table.com/control-chars.php#type) | **Description in C0 of ISO 646** |
| [Ctrl] @ | 0 | 00 | **NUL** | Null |  | A control character used to accomplish media-fill or time-fill. Null characters may be inserted into or removed from a stream of data without affecting the information content of that stream. But then the addition or removal of these characters may affect the information layout and/or the control of equipment. |
| [Ctrl] A | 1 | 01 | **STX** | Start of Header | TC | A transmission control character used as the first character of a heading of an information message. |
| [Ctrl] B | 2 | 02 | **SOT** | Start of Text | TC | A transmission control character which precedes a text and which is used to terminate a heading. |
| [Ctrl] C | 3 | 03 | **ETX** | End of Text | TC | A transmission control character which terminates a text. |
| [Ctrl] D | 4 | 04 | **EOT** | End of Transmission | TC | A transmission control character used to indicate the conclusion of the transmission of one or more texts.. |
| [Ctrl] E | 5 | 05 | **ENQ** | Enquiry | TC | A transmission control character used as a request for a response from a remote station; the response may include station identification and/or station status. When a "Who are you" function is required on the general switched transmission network, the first use of ENQ after the connection is established shall have the meaning "Who are you" (station identification). Subsequent use of ENQ may, or may not, include the function "Who are you", as determined by agreement. |
| [Ctrl] F | 6 | 06 | **ACK** | Acknowledge | TC | A transmission control character transmitted by a receiver as an affirmative response to the sender. |
| [Ctrl] G | 7 | 07 | **BEL** | Bell |  | A control character that is used when there is a need to call for attention; it may control alarm or attention devices. |
| [Ctrl] H | 8 | 08 | **BS** | BackSpace | FE | A format effector which moves the active position one character position backwards on the same line. |
| [Ctrl] I | 9 | 09 | **HT** | Horizontal Tabulation | FE | A format effector which advances the active position to the next pre-determined character position on the same line. |
| [Ctrl] J | 10 | 0A | **LF** | Line Feed | FE | A format effector which advances the active position to the same character position of the next line. |
| [Ctrl] K | 11 | 0B | **VT** | Vertical Tabulation | FE | A format effector which advances the active position to the same character position on the next pre-determined line. |
| [Ctrl] L | 12 | 0C | **FF** | Form Feed | FE | A format effector which advances the active position to the same character position on a pre-determined line of the next form or page. |
| [Ctrl] M | 13 | 0D | **CR** | Carriage Return | FE | A format effector which moves the active position to the first character position on the same line. |
| [Ctrl] N | 14 | 0E | **SO** | Shift Out |  | A control character which is used in conjunction with SHIFT IN and ESCAPE to extend the graphic character set of the code. It may alter the meaning of octets 33 - 126 (dec.). The effect of this character when using code extension techniques is described in International Standard ISO 2022. |
| [Ctrl] O | 15 | 0F | **SI** | Shift In |  | A control character which is used in conjunction with SHIFT OUT and ESCAPE to extend the graphic character set of the code. It may reinstate the standard meanings of the octets which follow it. The effect of this character when using code extension techniques is described in International Standard ISO 2022. |
| [Ctrl] P | 16 | 10 | **DLE** | Data Link Escape | TC | A transmission control character which will change the meaning of a limited number of contiguously following characters. Its is used exclusively to provide supplementary data transmission control functions. Only graphic characters and transmission control characters can be used in DLE sequences. |
| [Ctrl] Q | 17 | 11 | **DC1** | Device Control 1(XON) |  | A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to restore a device to the basic mode of operation (see also DC2 and DC3), or for any other device control function not provided by other DCs. |
| [Ctrl] R | 18 | 12 | **DC2** | Device Control 2 |  | A device control character which is primarily intended for turning on or starting an ancillary device. If it is not required for this purpose, it may be used to set a device to a special mode of operation (in which case DC1 is used to restore normal operation), or for any other device control function not provided by other DCs. |
| [Ctrl] S | 19 | 13 | **DC3** | Device Control 3(XOFF) |  | A device control character which is primarily intended for turning off or stopping an ancillary device. This function may be a secondary level stop, for example, wait, pause, stand-by or halt (in which case DC1 is used to restore normal operation). If it is not required for this purpose, it may be used for any other device control function not provided by other DCs. |
| [Ctrl] T | 20 | 14 | **DC4** | Device Control 4 |  | A device control character which is primarily intended for turning off, stopping or interrupting an ancillary device. If it is not required for this purpose, it may be used for any other device control function not provided by other DCs. |
| [Ctrl] U | 21 | 15 | **NAK** | Negative acknowledge | TC | A transmission control character transmitted by a receiver as a negative response to the sender. |
| [Ctrl] V | 22 | 16 | **SYN** | Synchronous Idle | TC | A transmission control character used by a synchronous transmission system in the absence of any other character (idle condition) to provide a signal from which synchronism may be achieved or retained between data terminal equipment. |
| [Ctrl] W | 23 | 17 | **ETB** | End of Transmission Block | TC | A transmission control character used to indicate the end of a transmission block of data where data is divided into such blocks for transmission purposes. |
| [Ctrl] X | 24 | 18 | **CAN** | Cancel |  | A character, or the first character of a sequence, indicating that the data preceding it is in error. As a result, this data is to be ignored. The specific meaning of this character must be defined for each application and/or between sender and recipient. |
| [Ctrl] Y | 25 | 19 | **EM** | End of Medium |  | A control character that may be used to identify the physical end of a medium, or the end of the used portion of a medium, or the end of the wanted portion of data recoreded on a medium. The position of this character does not necessarily correspond to the physical end of the medium. |
| [Ctrl] Z | 26 | 1A | **SUB** | Substitute |  | A control character used in the place of a character that has been found to be invalid or in error. SUB is intended to be introduced by automatic means. |
| [Ctrl] [ | 27 | 1B | **ESC** | Escape |  | A control character which is used to provide additional control functions. It alters the meaning of a limited number of contiguously following bit combinations. The use of this character is specified in International Standard ISO 2022. |
| [Ctrl] \ | 28 | 1C | **FS** | File Separator | IS | A control character used to separate and qualify data logically; its specific meaning has to be specified for each application. If this character is used in hierarchical order, it delimits a data item called a file |
| [Ctrl] ] | 29 | 1D | **GS** | Group Separator | IS | A control character used to separate and qualify data logically; its specific meaning has to be specified for each application. If this character is used in hierarchical order, it delimits a data item called a *group*. |
| [Ctrl] ^ | 30 | 1E | **RS** | Record Separator | IS | A control character used to separate and qualify data logically; its specific meaning has to be specified for each application. If this character is used in hierarchical order, it delimits a data item called a *record*. |
| [Ctrl] \_ | 31 | 1F | **US** | Unit Separator | IS | A control character used to separate and qualify data logically; its specific meaning has to be specified for each application. If this character is used in hierarchical order, it delimits a data item called a *unit*. |
|  | 127 | 7F | **DEL** | Delete |  | Delete |

|  |  |
| --- | --- |
| **Type** | **Description** |
| **FE** | Format Effector |
| **IS** | Information Separator |
| **TC** | Transmission Control |

Notes:

On a keyboard, it is often possible to generate a control code using the control (Ctrl, Ctl) key and a normal key.  
The column "**Keys**" shows the keys used to generate the control code.

说明：

七位ASCII定义了33个代码作为控制字符，它们是0到31、以及127，（位于0x00-0x1F及0x7F）。

兼容的八位ISO/IEC 8859-1加上了从ISO/IEC 6429定义的从128到159的32个代码，位于0x80-0x9F。

在python中，可以使用如下代码过滤字符串中的控制字符：

def strip\_control\_characters(str\_input):

 if str\_input:

 import re

 # unicode invalid characters

 RE\_XML\_ILLEGAL = u'([\u0000-\u0008\u000b-\u000c\u000e-\u001f\ufffe-\uffff])' + \

   u'|' + \

   u'([%s-%s][^%s-%s])|([^%s-%s][%s-%s])|([%s-%s]$)|(^[%s-%s])' % \

   (unichr(0xd800),unichr(0xdbff),unichr(0xdc00),unichr(0xdfff),

    unichr(0xd800),unichr(0xdbff),unichr(0xdc00),unichr(0xdfff),

    unichr(0xd800),unichr(0xdbff),unichr(0xdc00),unichr(0xdfff),

    )

 str\_input = re.sub(RE\_XML\_ILLEGAL, "", input)

 # ascii control characters

 str\_input = re.sub(r"[\x01-\x1F\x7F]", "", input)

 return str\_input

http://www.jb51.net/article/104804.htm