Tech Info Library

ASCII Control Characters

This article last reviewed: 1 November 1988

Code sets provide a means of electronically expressing and communicating
characters. There are three types of characters:

1. Alphabetic characters are the letters of the aphabet
   and special characters like <, *, @, ?, ", etc.

2. Numeric characters are numbers.

3. Control characters are unique characters that
   communicate specific functions (e.g., form feed,
   carriage return, escape, etc.) to a receiving unit.

Each character in a code set is expressed by a unique binary code. Morse code
is an example of a code set. Each Morse character has a unique combination of
dots and dashes that equate to the ones and zeros in a binary code.

Most people understand alphabetic and numeric characters. However, many
people are not as familiar with control characters, so the following
discussion will explain the control characters in more detail.

ASCII Control characters are used for five main functions:

1. COMMUNICATION CONTROL FUNCTIONS - SOH, STX, ETX, EOT, ENQ, ACK, NAK,
   DLE, SYN, ETB

2. CODE EXTENSION FUNCTIONS - Shift Out, Shift In, Escape

3. FORMAT EFFECTOR FUNCTIONS - Backspace, Tabs, Line Feed, Form Feed,
   Carriage Return

4. INFORMATION SEPARATOR FUNCTIONS - File Separator, Group Separator,
   Record Separator, Unit Separator

5. APPLICATION-DEFINED FUNCTIONS - Device Control 1 through 4, Bell

1. Communication Control Functions
These characters are mainly used in synchronous transmission character
oriented protocols (such as Bisynchronous).

Since data communication between personal computer systems is usually
asynchronous, we won't discuss the function of these characters here.

2. Code Extension Function
Code extension (ce) characters such as escape <ESC> are embedded in text to
enable devices to recognize device-specific control codes not included in
the code set.

The most commonly used ce character is <ESC>. It notifies a receiving unit
that the character or characters following it are to be interpreted as a
control code. A sequence of characters that starts with an <ESC> and makes
up a control code is popularly called an "escape sequence."

For example, it might be desirable that a printer change the size of
characters in the middle of a printing job. Since this function is not
common to all types of equipment, an ASCII control character has not been
assigned for it. Therefore many printer manufacturers have had to create
escape sequences for this purpose.

When a printer receives and recognizes the ESC character, it interprets the
following characters (up to the next RETURN character) as a control code, and
(in our example) changes its print size accordingly.

Escape sequences are manufacturer-specific, so different devices may have
different control codes to accomplish the same function (although there are
ANSI standards which many conform to).

3. Format Effector Function
Format effector (FE) characters (for instance, carriage returns, line feeds)
are embedded in the text sent to output devices like printers, CRT terminals,
etc. These characters give these devices instructions on how to physically
format the text when it is output.

4. Information Separator Function
Information separator (IS) characters allow the receiving unit to group the
received data into a database hierarchy (file to group to record to unit,
etc.).

5. Application-Defined Function
After all of the desired characters for a code set have been defined, there
may be several codes (unique combinations of bits) left over. These codes
are available for whatever purpose the designer may wish to assign them.

For example, in ASCII the DC1 and DC3 (Device Control) application defined
codes are often used to start and stop data in the XON/XOFF flow control
protocol (for more info on flow control protocols search on HTS and flow
control).
Copyright 1988 Apple Computer, Inc.

Keywords: HTS,CNFG

===

This information is from the Apple Technical Information Library.

19960215 11:05:19.00

Tech Info Library Article Number: 1373