ISO Latin 9 as compared with ISO Latin 1

The ISO Latin 9 (ISO 8859-15) character set differs from the well-known ISO Latin 1 (ISO 8859-1) character set in a few positions only. The euro sign and some national letters used e.g. in French and Finnish have been introduced and some rarely used special characters omitted.

ISO Latin 9 is a recent addition to the ISO 8859 family of character sets, published as a standard 1999-03-15. The proposal (also available in French) was in the final voting phase (during the period 1998-01-10/1998-12-01) within ISO; it acquired the required support, so it was published as a standard without modifications.

ISO Latin 9 differs from the widely used ISO Latin 1 (ISO 8859-1) in some character positions only. When viewed as a modification of ISO Latin 1, ISO Latin 9 replaces some rarely needed characters by euro sign - a politically important symbol for the currency of the European Monetary Union (EMU) - and the letters oe ligature and capital Y with diaeresis (used in French) and letters s with caron and z with caron (used in Finnish and other languages). The following table presents the differences in detail.

Code position			**ISO Latin 1**			**ISO Latin 9**		
dec	**oct**	**hex**	**glyph**	**U+*nnnn***	**Unicode name**	**glyph**	**U+*nnnn***	**Unicode name**
164	244	A4	¤	U+00A4	currency symbol	€	U+20AC	euro sign (U+FB01)
166	246	A6	¦	U+00A6	broken bar	Š	U+0160	latin capital letter s with caron
168	250	A8	¨	U+00A8	diaeresis	š	U+0161	latin small letter s with caron
180	264	B4	´	U+00B4	acute accent	Ž	U+017D	latin capital letter z with caron
184	270	B8	¸	U+00B8	cedilla	ž	U+017E	latin small letter z with caron
188	274	BC	¼	U+00BC	vulgar fraction one quarter	Œ	U+0152	latin capital ligature oe
189	275	BD	½	U+00BD	vulgar fraction one half	œ	U+0153	latin small ligature oe
190	276	BE	¾	U+00BE	vulgar fraction three quarters	Ÿ	U+0178	latin capital letter y with diaeresis

General notes

Originally the project which lead to the creation of ISO Latin 9 used the working name "Latin Alphabet Number Zero" for it. Therefore it has often been referred to as "Latin 0".

The famous *ISO 8859 Alphabet Soup* document contains a complete ISO Latin 9 code table.

Within the European Union (EU), the euro currency has great political and practical importance. The euro sign is both a character that will be used widely and a symbol of the monetary union. Therefore, within EU, and in countries applying for EU membership, it will become a crucial requirement that euro sign is available. For technical information on this, see section *Euro symbol* in *IT impact of the Euro*.

ISO Latin 9 contains both euro sign and some important national letters, and the ISO Latin 1 characters replaced by them are *very* rarely used, as the document *The ISO Latin 1 character repertoire - a description with usage notes* explains in some detail. Since eight-bit character encodings are still very important, it seems more than likely that ISO Latin 1 will be generally replaced by ISO Latin 9, within Europe at least.

Some of the ISO Latin 9 characters which do not belong to ISO Latin 1 appear in the so-called Windows character set (WinLatin1) as originally defined. These characters are s with caron, oe ligature, and capital Y with diaeresis. Moreover, the euro sign has been added (using code position 80 hexadecimal) into new variants of the Windows character sets. This means that glyphs for the characters are widely available on PCs, although an update may be needed for the euro sign, but **the use of the characters requires solutions to character encoding problems**. Compare with the document *On the use of some MS Windows characters in HTML*, which explains the problems caused by differences between ISO Latin 1 and the MS Windows character set. Using ISO Latin 9, *some* of these problems can be solved.

However, **changing from ISO Latin 1 to ISO Latin 9** is a complicated task, partly because so many program assume ISO Latin 1 as the default character set. See section *Information about encoding* in *A tutorial on character code issues* for the big picture. In particular, a program which sends ISO Latin 9 encoded text data over the Internet, e.g. by E-mail or using the HTTP protocol, should specify `Content-Type: text/plain;charset=iso-8859-15`.

Notes on the characters

The euro sign, denoting the currency of the European Monetary Union (EMU), was added to Unicode, in version 2.1 as U+20AC. Notice that it is distinct from the euro-currency sign (U+20A0), which is of minor historical interest only. The **presentation of the euro sign is problematic** especially in contexts where data needs to be transmitted over networks or otherwise from one data processing system to another, e.g. on Web pages, on Usenet, and E-mail. Generally, in such situations it is still preferable to use the **word** "euro", in applicable language, orthography, and form. See *Spelling of the words "euro" and "cent" in the official Community languages* (also available in PDF format, where the Greek words are presented properly). In special situations, like international banking business, where ISO 4217 codes are used for currencies, the abbreviation to be used is EUR. As mentioned above, there is a lot of technical information on the euro symbol in the EU documentation *IT impact of the Euro*.

Letters "s" and "z" with caron are used in the official orthography of the Finnish language; see an official statement on this (available in Finnish, too) . (In practice, they are quite often replaced by character pairs "sh" and "zh", partly due to character code and font problems.) They are also used e.g. in Estonian and Czech as well as in the official international transliteration (Latinization) scheme for the Cyrillic alphabet, ISO 9.

The ligature of "o" and "e" is considered as a separate character in French. It is also used e.g. in one style of writing Latin (cf. to the use of "the letter ae" as a ligature of "a" and "e", e.g. writing "Caesar" as "Cæsar").

The latin capital letter y with diaeresis is used in some French names when they are written in capital letters.

This document exists in two formats:

- HTML document, at `http://www.hut.fi/u/jkorpela/latin9.html`
- PostScript document (generated from the HTML document using html2ps), at `http://www.hut.fi/u/jkorpela/latin9.ps`

Date of last update: 1999-09-20

Jukka Korpela, Jukka.Korpela@hut.fi