Language Environment Code Set Conversion

Language Environment for VSE/ESA (LE/VSE) provides code set converters for direct conversion and indirect conversion via Universal Character Set 2 (UCS-2).

- The code set converters supplied for direct conversion are named EDCUaabb, where aa is the CC code of the source code set and bb is the CC code of the destination code set.

- The UCS-2 converters are named EDCUUccU, where cc is the CC code of the used code set. For indirect conversation via UCS-2 two converters are used.

LE/VSE also supports conversion from or to UCS-2 and UTF-8 (USC Transformation Format 8) using the UCS-2 converters. See LE/VSE C Run-Time Programming Guide (SC33-6688) for description of the supplied converters.

These converters can be used with:

1. The C Run-Time functions iconv_open(), iconv(), iconv_close() to convert a sequence of characters from one encoded character set into a sequence of corresponding characters in another encoded character set. See LE/VSE C Run-Time Library Reference (SC33-6989) for descriptions of the iconv functions.

2. The iconv utility to convert a file from one code set encoding to another.

In the following iconv utility example, the input file INPUT.FILE (in ASCII code page IBM-850) is converted to the output file OUTPUT.FILE (in EBCDIC code page IBM-1140).

```
// JOB ICNVSAMP
// LIBDEF *,SEARCH=PRD2.SCEEBASE
// EXEC EDCICONV,PARM='IFILE(''INPUT.FILE''),OFILE(''OUTPUT.FILE''),   X
                  FROMCODE(IBM-850),TOCODE(IBM-1140)'
/*
// EXEC LISTLOG
/&
```

See LE/VSE C Run-Time Programming Guide (SC33-6688) for descriptions of the iconv utility.

Available Unicode codesets:

Codeset Name	CC code	Comment
UCS-2	U2	Universal Character Set 2
UTF-8	F8	USC Transformation Format 8

Available EBCDIC codesets:

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-500	EO	NO	NO	International #5. Euro supported code page is 1148; ICS is replaced by euro at X'9F'.
IBM-1148	HO	NO	YES	International ECECP. This code page is based on code page 500 with ICS replaced by euro.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-1047	EY	NO	NO	Latin 1/Open Systems. This code page is meant for limited use to meet specific requirements. Euro supported code page is 924; ICS is replaced by euro at X'9F'.
IBM-924	EZ	YES	YES	Latin 9 EBCDIC. This code page is the EBCDIC version of Latin 9. It is based on code page 1047 with eight characters replaced for Latin 9 and two synonym characters replaced.
IBM-037	EA	NO	NO	USA/Canada - CECP. Euro supported version of this code page is 1140; ICS is replaced by euro at X'9F'.
IBM-28709	EA	NO	NO	Alias or subset of IBM-037.)
IBM-1140	HA	NO	YES	USA, Canada, etc. ECECP. This code page is based on code page 037 with ICS replaced by euro at X'9F'.
IBM-273	EB	NO	NO	Germany F.R./Austria - CECP. Euro supported code page is 1141; ICS is replaced by euro at X'9F'.
IBM-1141	HB	NO	YES	Austria, Germany ECECP. This code page is based on code page 273 with ICS replaced by euro at X'9F'.
IBM-274	EC	NO	NO	Old Belgium Code Page.
IBM-278	EF	NO	NO	Finland, Sweden - CECP. Euro supported code page is 1143; ICS is replaced by euro at X'5A'.
IBM-1143	HF	NO	YES	Finland, Sweden ECECP. This code page is based on code page 278 with ICS replaced by euro.
IBM-280	EG	NO	NO	Italy - CECP. Euro supported code page is 1144; ICS is replaced by euro at X'9F'.
IBM-1144	HG	NO	YES	Italy ECECP. This code page is based on code page 280 with ICS replaced by euro.
IBM-282	EI	NO	NO	Portugal - CECP.
IBM-284	EJ	NO	NO	Spain/Latin America - CECP. Euro supported code page is 1145; ICS is replaced by euro at X'9F'.
IBM-1145	HJ	NO	YES	Spain, Latin America (Spanish) ECECP. This code page is based on code page 284 with ICS replaced by euro.
IBM-297	EM	NO	NO	France - CECP. Euro supported code page is 1147; ICS is replaced by euro at X'9F'.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-1147	HM	NO	YES	France ECECP. This code page is based on code page 297 with ICS replaced by euro.
IBM-875	ES	NO	NO	Greece. Euro supported CCSID of this codepage is 4971; with euro symbol added to X'FC'.
IBM-4971	HS	YES	YES	Greece. This code page is based on code page 875 with euro symbol added at X'FC'.
IBM-277	EE	NO	NO	Denmark, Norway - CECP. Euro supported code page is 1142; ICS is replaced by euro at X'5A'.
IBM-1142	HE	NO	YES	Denmark, Norway ECECP. This code page is based on code page 277 with ICS replaced by euro.
IBM-285	EK	NO	NO	United Kingdom - CECP. Euro supported code page is 1146; ICS is replaced by euro at X'9F'. Direct converters from and to IBM-1047 and IBM-924 only.
IBM-1146	HK	NO	YES	UK ECECP. This code page is based on code page 285 with ICS replaced by euro.
IBM-870	EQ	NO	NO	Latin 2 - EBCDIC Multilingual. Euro supported version of this code page is 1153; ICS is replaced with euro symbol at X'9F'.
IBM-1153	MB	YES	YES	EBCDIC Latin 2 Multilingual with euro. This code page is based on CP 870 with ICS at X'9F' replaced by euro.
IBM-1026	EW	NO	NO	Latin #5 - Turkey. Euro supported version of this code page is 1155; ICS is replaced by euro symbol at X'9F'.
IBM-1155	HW	YES	YES	EBCDIC Turkey with euro. This code page is based on CP 1026 with ICS at X'9F' replaced by euro.
IBM-871	ER	NO	NO	Iceland - CECP. Euro supported code page is 1149; ICS is replaced by euro at X'9F'.
IBM-1149	HR	NO	YES	Iceland ECECP. This code page is based on code page 871 with ICS replaced by euro.
IBM-1122	FD	NO	NO	Estonia, EBCDIC. Euro supported version of this codepage is 1157; ICS is replaced by euro symbol at X'5A'.
IBM-1157	HD	YES	YES	EBCDIC Estonia with euro. This code page is based on CP 1122 with ICS at X'5A' replaced by euro.
IBM-880	ET	NO	NO	Cyrillic, Multilingual.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-1025	FE	NO	NO	Cyrillic, Multilingual. Euro supported version of this codepage is 1154; SM24 is replaced by euro symbol at X'E1'.
IBM-1154	HT	YES	YES	EBCDIC Cyrillic, Multilingual with euro. This code page is based on CP 1025 with SM24 at X'E1' replaced by euro.
IBM-1112	GD	NO	NO	Baltic - Multilingual, EBCDIC. Euro supported version of this codepage is 1156; ICS is replaced by euro symbol at X'9F'.
IBM-1156	HZ	YES	YES	EBCDIC Baltic Multi with euro. This code page is based on CP 1112 with ICS at X'9F' replaced by euro.
IBM-424	FB	NO	NO	Israel (Hebrew). Euro supported code page is 12712.
IBM-12712	HH	YES	YES	Israel (Hebrew). This code page is based on code page 424 with euro at \x9C and new sheqel.
IBM-420	FF	NO	NO	Arabic Bilingual. Euro supported CCSID of this codepage is 16804; with euro symbol added to X'FA'.
IBM-16804	HC	YES	YES	Arabic Bilingual with euro. This code page is based on code page 420 with euro symbol added at X'FA'.
IBM-275	ED	NO	NO	Brazil - CECP.
IBM-838	EP	NO	NO	Thai with Low Tone Marks & Ancient Characters. Euro supported version of this code page is 1160. BZ500300 is replaced by euro at X'FE'.
IBM-1160	HP	YES	YES	Thai with Low Tone Marks & Ancient Characters. This code page is based on code page 838 with BZ500300 at X'FE' replaced with euro.
IBM-281	EH	NO	NO	Japan (Latin) - CECP. SBCS. Direct converters from and to IBM-1047 and IBM-1148 only.
IBM-1027	EX	NO	NO	Japanese (Latin) Extended. SBCS. Euro supported code page is 5123. Euro currency symbol is added at X'E1'.
IBM-5123	HX	YES	YES	Japanese (Latin) Extended. SBCS. This code page is based on code page 1027 with euro symbol added at X'E1'.
IBM-290	EL	NO	NO	Japanese (Katakana) Extended. SBCS. Euro supported code page is 8482; Euro currency symbol is added at X'E1'.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-8482	HL	YES	YES	Japanese (Katakana) Extended. SBCS. This code page is based on code page 290 with euro symbol added at X'E1'.
IBM-300	EN	NO	NO	Japanese DBCS.
IBM-4396	EN	NO	NO	Subset of IBM-300. DBCS.
IBM-939	EV	NO	NO	Japanese (Latin) Extended. SBCS IBM-1027 combined with DBCS IBM-300.
IBM-5035	EV	NO	NO	Subset of IBM-939.
IBM-930	EU	NO	NO	Japanese (Katakana) Extended. SBCS IBM-290 combined with DBCS IBM-300.
IBM-5026	EU	NO	NO	Subset of IBM-930. SBCS IBM-290 combined with DBCS IBM-4396.
IBM-833	GP	NO	NO	Korean Extended. SBCS.
IBM-834	GQ	NO	NO	Korean DBCS.
IBM-933	GZ	NO	NO	Korean Extended. SBCS IBM-833 combined with DBCS IBM-834.
IBM-1364	GU	NO	YES	Korean Extended. SBCS IBM-13121 combined with DBCS IBM-4930.
IBM-835	GO	NO	NO	Taiwan, Traditional Chinese. DBCS.
IBM-937	GW	NO	NO	Taiwan, Traditional Chinese. SBCS IBM-037 combined with DBCS IBM-835.
IBM-1371	KA	YES	YES	Taiwan, Traditional Chinese. SBCS IBM-1159 combined with DBCS IBM-9027.
IBM-836	GL	NO	NO	PRC, Simplified Chinese Extended. SBCS.
IBM-837	GM	NO	NO	PRC, Simplified Chinese. DBCS.
IBM-935	GY	NO	NO	PRC, Simplified Chinese Extended. SBCS IBM-836 combined with DBCS IBM-837.
IBM-5031	GY	NO	NO	Alias or subset of IBM-935.

Available ASCII codesets:

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-437	AV	YES	NO	Personal Computer.
IBM-850	AA	NO	NO	Personal Computer - Multilingual Page. Euro supported code page is 858; Dotless i is replaced by euro at X'D5'.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-4946	AA	NO	NO	Alias or subset of IBM-850.
IBM-858	AI	YES	YES	Personal Computer - Multilingual with euro. This code page is identical to code page 850 except that X'D5' has been changed from "dotless i" to "euro symbol". Direct converters from and to IBM-1047, IBM-1140, IBM-1141, IBM-1142, IBM-1143, IBM-1144, IBM-1145, IBM-1146, IBM-1147, IBM-1148, and IBM-1149 only.
IBM-859	AK	YES	YES	PC Latin 9. This code page is based on code page 850 but contains the Latin 9 character set. The dotless i is replaced with euro (as in CP 858).
IBM-819	I1	NO	NO	ISO/ANSI Multilingual. This code page meets the requirements of the external standard ISO 8859; Part 1.
ISO8859-1	I1	NO	NO	Alias of IBM-819.
IBM-1252	DA	NO	NO	Windows, Latin 1. This code page is a superset of ISO 8859-1. Euro supported CCSID of this codepage is 5348; with euro symbol added to X'80'.
IBM-5348	DJ	YES	YES	Windows, Latin 1 with euro. This code page is based on code page 1252 with euro symbol added at X'80'.
IBM-912	I2	NO	NO	Latin 2 - ISO. This code page meets the requirements of the external standard ISO 8859; Part 2.
ISO8859-2	I2	NO	NO	Alias of IBM-912.
IBM-852	CB	NO	NO	Latin 2 - Personal Computer. Euro supported version of this code page is 9044; Euro currency symbol is added at X'AA'.
IBM-9044	NG	YES	YES	Latin 2 - Personal Computer with euro. This code page is based on code page 852 with euro symbol added at X'AA'.
IBM-1250	DB	NO	YES	Windows, Latin 2. The characters in this code page are a superset of the characters of ISO 8859-2. 15 common characters are at different code points; the remaining common characters are at the same code points.
IBM-914	I4	NO	NO	Latin 4. This code page meets the requirements of the external standard ISO 8859; Part 4.
ISO8859-4	I4	NO	NO	Alias of IBM-914.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-915	I5	NO	NO	Cyrillic, 8-Bit. This code page meets the requirement of the external standard ISO 8859; Part 5.
ISO8859-5	I5	NO	NO	Alias of IBM-915.
IBM-855	CE	NO	NO	Cyrillic - Personal Computer. Euro supported version of this codepage is 872; ICS is replaced by euro symbol at X'CF'.
IBM-872	LE	YES	YES	Cyrillic - PC with euro. This code page is based on CP 855 with ICS at X'CF' replaced by euro.
IBM-866	BE	NO	NO	PC Data, Cyrillic, Russian. Euro supported version of this codepage is 808; ICS is replaced by euro symbol at X'FD'.
IBM-808	LF	YES	YES	PC Data, Cyrillic, Russian with euro. This code page is based on 866; with ICS replaced by euro at X'FD'.
IBM-1124	AU	YES	NO	Cyrillic, Ukraine.
IBM-1251	DE	NO	YES	Windows, Cyrillic. The characters in this code page are a superset of the characters of ISO 8859-5 (CP 00915). Many characters are not at the same code points.
IBM-1089	I6	NO	NO	Arabic Code Page, Data Storage & Interchange. This code page conforms to ISO 8859-6.
ISO8859-6	I6	NO	NO	Alias of IBM-1089.
IBM-1046	AF	NO	NO	Arabic Extended. This code page meets the character set requirements of ISO 8859/6. Euro supported CCSID of this codepage is 9238; with euro symbol added at X'FF'.
IBM-9238	LI	YES	YES	Arabic Extended with euro. This code page is based on code page 1046 with euro symbol added at X'FF'.
IBM-864	CF	NO	NO	Arabic - Personal Computer. Euro supported CCSID of this codepage is 17248; with euro symbol added to X'A7'.
IBM-17248	NJ	YES	YES	Arabic - Personal Computer with euro. This code page is based on code page 864 with euro symbol added at X'A7'.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-1256	DF	NO	NO	Windows, Arabic. The characters in this code page are a superset of the characters of ISO 8859-6. 25 common characters are at different code points; the remaining common characters are at the same code points. Note however that there are 70 additional characters in CP 01256. Euro supported CCSID of this codepage is 5352; with euro symbol added to X'80'.
IBM-5352	NF	YES	YES	Windows, Arabic with euro. This code page is based on code page 1256 with euro symbol added at X'80'.
IBM-813	I7	NO	NO	Greece ISO 8859-7. This code page meets the requirements of the external standard ISO 8859; Part 7. Euro supported CCSID of this codepage is 4909; with euro symbol added to X'A4'.
ISO8859-7	I7	NO	NO	Alias of IBM-813.
IBM-4909	IA	YES	YES	Greece ISO 8859-7 with euro. This code page is based on code page 813 with euro symbol added at X'A4'.
IBM-869	CG	NO	NO	Greece - Personal Computer. Euro supported CCSID of this codepage is 9061; with euro symbol added to X'87'.
IBM-9061	LG	YES	YES	Greece - Personal Computer with euro. This code page is based on code page 869 with euro symbol added at X'87'.
IBM-1253	DG	NO	NO	Windows, Greek. The characters in this code page are a superset of the characters of ISO 8859-7. 4 common characters are at different code points; the remaining common characters are at the same code points. Euro supported CCSID is 5349; with euro character addition at X'80'.
IBM-5349	DK	YES	YES	Windows, Greek with euro.
IBM-916	I8	NO	NO	Hebrew (Latin). This code page meets the requirements of the external standard ISO 8859; Part 8.
ISO8859-8	I8	NO	NO	Alias of IBM-916.
IBM-862	BH	NO	NO	Israel - Personal Computer. A new code page 867 was made in which SP03 have been replaced by euro at X'AD'; new sheqel and direction controls.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-867	LJ	YES	YES	Israel - Personal Computer. This code page is based on code page 862 except SP03 have been replaced by euro at X'AD'; new sheqel and direction controls.
IBM-856	CH	NO	NO	Hebrew - Personal Computer. (This code page was revised in April 1998 to include euro; new sheqel and direction controls.)
IBM-1255	DH	NO	NO	Windows, Hebrew. The characters in this code page are a superset of the characters of ISO 8859-8 (CP 00916) except for new sheqel which replaces ICS; and divide symbol, multiply symbol, cedilla, and double underscore which are missing. 1 common character (SM57) is at a different code point; the remaining common characters are at the same code points. Euro supported CCSID of this codepage is 5351; with euro symbol added to X'80'.
IBM-5351	NH	YES	YES	Windows, Hebrew with euro. This code page is based on code page 1255 with euro symbol added at X'80'.
IBM-920	I9	NO	NO	Latin #5 - Turkey. This code page meets the requirements of the external standard ISO 8859; Part 9.
ISO8859-9	I9	NO	NO	Alias of IBM-920.
IBM-5350	NI	YES	YES	Windows, Turkish with euro. This code page is a superset of ISO 8859-9. This code page is based on code page 1254 with euro symbol added at X'80'.
IBM-922	AD	NO	NO	Estonia, similar to ISO 8859-x. Euro supported version of this codepage is 902; ICS is replaced by euro symbol at X'A4'.
IBM-902	LD	YES	YES	8-bit Estonia with euro. This code page is based on CP 922 with ICS at X'A4' replaced by euro.
IBM-921	BD	NO	NO	Baltic - Multilingual, superset of ISO 8859-13. Euro supported version of this codepage is 901; ICS is replaced with euro symbol at X'A4'.
IBM-901	LH	YES	YES	PC Baltic Multi with euro. This code page is based on CP 921 with ICS at X'A4' replaced by euro.
IBM-861	CA	NO	NO	Iceland - Personal Computer.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-874	BU	NO	NO	Thai with Low Tone Marks & Ancient Chars - PC. This code page is almost a superset of ISO-8859-11. Euro supported version of this code page is 1161. BZ500300 is replaced by euro at X'DE'.
TIS-620	BU	NO	NO	Alias or subset of IBM-874.
IBM-1161	LU	YES	YES	Thai with Low Tone Marks & Ancient Chars - PC. This code page is almost a superset of ISO-8859-11. This code page is based on code page 874 with BZ500300 at X'DE' replaced with euro.
IBM-942	AB	NO	NO	Japanese PC Code Page. SBCS IBM-1041 combined with DBCS IBM-301.
IBM-932	AB	NO	NO	Subset of IBM-942. SBCS IBM-897 combined with DBCS IBM-301.
IBM-301	AB	NO	NO	Subset of IBM-942. DBCS.
IBM33722	AC	NO	NO	Japanese EUC Code Page. SBCS IBM-895/IBM-896 combined with DBCS IBM-952/IBM-953.
IBM-eucJP	AC	NO	NO	Alias of IBM33722.
IBM-1088	CP	NO	NO	Korean PC Code Page. SBCS.
IBM-951	CQ	NO	NO	Korean PC Code Page. DBCS.
IBM-949	CZ	NO	NO	Korean PC Code Page. SBCS IBM-1088 combined with DBCS IBM-951.
IBM-1363	CU	NO	YES	Korean PC Code Page. SBCS IBM-1126 combined with DBCS IBM-1362.
IBM-970	BZ	NO	NO	Korean EUC Code Page. SBCS IBM-367 combined with DBCS IBM-971.
IBM-eucKR	BZ	NO	NO	Alias of IBM-970.
eucKR	BZ	NO	NO	Alias of IBM-970.
IBM-904	CN	NO	NO	Taiwan, Traditional Chinese PC Code Page. SBCS.
IBM-950	DW	NO	NO	Taiwan, Traditional Chinese PC Code Page. SBCS IBM-1114 combined with DBCS IBM-947.
BIG5	DW	YES	NO	Alias of IBM-950.
IBM-947	DW	NO	NO	Subset of IBM-950. DBCS.
IBM-1370	LA	YES	YES	Taiwan, Traditional Chinese PC Code Page. SBCS IBM-5210 combined with DBCS IBM-21427.

Codeset Name	CC code	New in z/VSE	Euro-Support	Comment
IBM-964	BW	NO	NO	Taiwan, Traditional Chinese EUC Code Page. SBCS IBM-367 combined with DBCS IBM-960/IBM-961.
IBM-eucTW	BW	NO	NO	Alias of IBM-964.
eucTW-1993	BW	NO	NO	Alias of IBM-964.
IBM-1115	CL	NO	NO	PRC, Simplified Chinese PC Code Page. SBCS.
IBM-1380	CM	NO	NO	PRC, Simplified Chinese PC Code Page. DBCS.
IBM-1381	CY	NO	NO	PRC, Simplified Chinese PC Code Page. SBCS IBM-1115 combined with DBCS IBM-1380.
IBM-1383	BY	NO	NO	PRC, Simplified Chinese EUC Code Page. SBCS IBM-367 combined with DBCS IBM-1382.
IBM-eucCN	BY	NO	NO	Alias of IBM-1383.

Trademarks

The following terms are trademarks of International Business Machines Corporation in the United States, or other countries, or both:

IBM, Language Environment, VSE/ESA, z/VSE

Microsoft, Windows, Windows NT and Windows logo, are trademarks or registered trademarks of Microsoft Corporation.

Other company, product, or service names, may be the trademarks or service marks of others.

Comments and Questions

Comments or questions on this documentation are welcome. Please send your comments to:

zvse@de.ibm.com