

Information processing; 8 bit character set for information interchange

Zusammenhang mit der von der International Organization for Standardization (ISO) herausgegebenen Internationalen Norm ISO 6937/2 – 1983 und den Internationalen Norm-Entwürfen ISO/DIS 4873 – 1984 und ISO/DIS 8859/1.2 – 1986, siehe Erläuterungen.

Inhalt

	Seite		Seite
1 Anwendungsbereich und Zweck	1	5.3 Multibyte Version des 8-Bit-Code mit diakritischen Zeichen (MBV8)	2
2 Bezeichnung	1	5.4 Erklärung der in den Code-Tabellen benutzten internationalen Kurzzeichen	6
3 Code-Tabellen	1	5.5 Erklärung der in den Code-Tabellen benutzten internationalen Kurzzeichen und Schriftzeichen	6
3.1 Allgemeiner Aufbau	1	Anhang A Identifikationsverfahren	16
3.2 Definition einer Version	2	Zitierte Normen	17
4 Allgemeine Grundsätze zur Bildung des 8-Bit-Code	2	Weitere Normen	17
5 Code-Versionen	2	Erläuterungen	18
5.1 Allgemeine Referenz-Version des 8-Bit-Code (ARV8)	2		
5.2 Deutsche Referenz-Version des 8-Bit-Code (DRV8)	2		

1 Anwendungsbereich und Zweck

Der in dieser Norm festgelegte 8-Bit-Code dient zur Übergabe bzw. Übermittlung sowie zur Ein- und Ausgabe von digitalen Daten in einer 8-Bit-Umgebung.

Eine Version dieses Code wird durch die Festlegung von Schrift- und Steuerzeichensätzen definiert. In dieser Norm werden nur die Schriftzeichensätze festgelegt. Die zur Bildung einer Version notwendigen Steuerzeichensätze werden abhängig von der Anwendung ausgewählt. Steuerzeichen sind in DIN 66 254*) festgelegt.

Die in 8 Bits codierten Zeichensätze sind aus dem in DIN 66 003 festgelegten, in 7 Bits codierten Zeichensatz abgeleitet und mit diesem verträglich. Außerdem stimmen ihre Merkmale mit den Code-Erweiterungsverfahren überein, die in DIN 66 203 beschrieben werden; darin werden neben anderen Dingen eine Familie von 8-Bit-Codes definiert.

Die vorliegende Norm ist auf solche Mitglieder der Familie von 8-Bit-Codes beschränkt, die als Grundmenge die Schriftzeichen einer Version des 7-Bit-Code nach DIN 66 003 haben.

Die Darstellung des Code auf Datenträgern ist in DIN 66 004 Teil 1 bis Teil 5 festgelegt.

2 Bezeichnung

Die Norm-Bezeichnung für die Deutsche Referenz-Version des 8-Bit-Code, DRV8, lautet:

Code DIN 66 303 – DRV8

Siehe auch Abschnitte 5.1, 5.2 und 5.3.

*) Z. Z. Entwurf

3 Code-Tabellen

3.1 Allgemeiner Aufbau

Durch eine Code-Tabelle wird sowohl ein Zeichensatz als auch die Codierung der Zeichen definiert. Die Code-Tabellen dieser Norm bestehen aus einer Anordnung von 16 Spalten und 16 Zeilen, die jeweils von 00 bis 15 benummert sind; dies ergibt 256 Plätze in einer Code-Tabelle.

Die einzelnen Zeichen des 8-Bit-Code können durch ihre Bitkombination in der Reihenfolge $b_8, b_7, b_6, b_5, b_4, b_3, b_2, b_1$ oder durch die Angabe von Spalten- und Zeilennummer ihres Platzes in der Code-Tabelle gekennzeichnet werden.

Beispiel:

Buchstabe F entspricht der Bitkombination 01000110 oder Platz 04/06

Die Gliederung der 8-Bit-Code-Tabelle und der Zusammenhang zwischen 7-Bit- und 8-Bit-Codes ist in DIN 66 203 angegeben.

Die Spalten 00 bis 07 enthalten Plätze für 128 Zeichen; die codierte Darstellung der Zeichen ist die gleiche wie im 7-Bit-Code, nur daß ein achttes, an höchster Position stehendes Bit hinzutritt, welches den Wert 0 hat.

Die Spalten 00 und 01 sind für einen Satz von 32 Steuerzeichen reserviert, den C0-Satz.

Die Spalten 02 bis 07 mit Ausnahme der Plätze 02/00 und 07/15 enthalten einen Satz von 94 Schriftzeichen, den G0-Satz. Der Platz 02/00 enthält das Leerzeichen (SP) und der Platz 07/15 ist nicht mit einem Schriftzeichen belegt.

Die Spalten 08 bis 15 enthalten Plätze für weitere 128 Zeichen; das achte Bit der codierten Darstellung dieser Zeichen hat den Wert 1.

Die Spalten 08 und 09 sind für einen Satz von 32 Steuerzeichen reserviert, den C1-Satz.

Die Spalten 10 bis 15 sind für einen Satz von 96 Schriftzeichen reserviert, den G1-Satz.

Fortsetzung Seite 2 bis 18

3.2 Definition einer Version

Eine zu verwendende Version wird durch ESC-Folgen definiert, wie in DIN 66 203 beschrieben, nämlich

- a) ESC 02/01 (F) für den Satz von Steuerzeichen (CO-Satz) in den Spalten 00 und 01
- b) ESC 02/08 (F) oder ESC 02/12 (F) für den Satz von Schriftzeichen (G0-Satz) in den Spalten 02 bis 07.
- c) ESC 02/02 (F) für den Satz von Steuerzeichen (C1-Satz) in den Spalten 08 und 09
- d) ESC 02/09 (F) oder ESC 02/13 (F) für den Satz von Schriftzeichen (G1-Satz) in den Spalten 10 bis 15.

Dabei steht (F) jeweils für eine der Bitkombinationen von 03/00 bis 07/14. Nähere Einzelheiten sind in DIN 66 203 angegeben.

4 Allgemeine Grundsätze zur Bildung des 8-Bit-Code

Die Codeerweiterungsregeln nach DIN 66 203 erlauben es, eine Vielzahl von verschiedenen strukturierten 8-Bit-Codes zu bilden. Für den Zweck dieser Norm erwies es sich als zweckmäßig, die eingeschränkten Möglichkeiten nach ISO/DIS 4873-1984 (DIN-Norm in Vorbereitung, siehe Erläuterungen) anzuwenden. Um eine größtmögliche Verträglichkeit mit dem 7-Bit-Code nach DIN 66 003 sicherzustellen, eine einheitliche Ausgangsbasis zu haben und ein Maximum an internationaler Kompatibilität zu erreichen, wurden für den allgemeinen Datenaustausch in den Spalten 02 bis 07 nur die in DIN 66 003 festgelegten Schriftzeichen des 7-Bit-Code vorgesehen. Ausnahme ist der Platz 02/04 in der Code-Tabelle 1 (ARV8), der mit \$ anstelle von ꝥ belegt wurde.

5 Code-Versionen

Der 8-Bit-Code ist in 3 Versionen festgelegt:

Code-Tabelle 1. Allgemeine Referenz-Version des 8-Bit-Code (ARV8)

Code-Tabelle 2. Deutsche Referenz-Version des 8-Bit-Code (DRV8)

Code-Tabelle 3. Multibyte Version des 8-Bit-Code mit diakritischen Zeichen (MBV8)

5.1 Allgemeine Referenz-Version des 8-Bit-Code (ARV8)

Bei der Allgemeinen Referenz-Version des 8-Bit-Code (Code-Tabelle 1) sind die Zeichen in den Spalten 02 bis 07 identisch mit den Zeichen in den Spalten 2 bis 7 der Internationalen Referenz-Version (IRV) des 7-Bit-Code nach DIN 66 003 mit Ausnahme des Platzes 02/04, der mit \$ anstelle von ꝥ belegt wurde.

Die Spalten 00 und 01 sowie 08 und 09 sind unbenutzt, solange sie nicht jeweils von einem genormten und mit entsprechenden ESC-Folgen definierten Steuerzeichensatz belegt werden.

Die Spalten 10 bis 15 sind mit Sonderzeichen sowie nationalen Buchstaben und Buchstaben mit diakritischen Zeichen oder Akzenten belegt, die in verschiedenen europäischen Sprachen verwendet werden. In diesen Spalten sind auch die Umlaute und der Buchstabe ß aufgenommen. Die Spalten 10 bis 15 entsprechen dem Lateinischen Alphabet Nummer 1 nach dem Internationalen Norm-Entwurf ISO/DIS 8859/1.2-1986. Jedes Schriftzeichen wird durch nur eine Bitkombination dargestellt. Eine Kombination der Schriftzeichen mit anderen Schriftzeichen ist nicht gestattet. Der Schriftzeichenvorrat ist gleich dem Schriftzeichensatz.

Diese Version wird eindeutig definiert durch die ESC-Folgen

CO-Satz: ESC 02/01 F
G0-Satz: ESC 02/08 04/02
C1-Satz: ESC 02/02 F
G1-Satz: ESC .../.../.../...¹⁾

Dabei steht (F) jeweils für eine der Bitkombinationen von 03/00 bis 07/14. Nähere Einzelheiten sind in DIN 66 203 angegeben.

Die Normbezeichnung für die Allgemeine Referenz-Version des 8-Bit-Code lautet:

Code DIN 66 303 – ARV8

5.2 Deutsche Referenz-Version des 8-Bit-Code (DRV8)

Diese Version wurde eingeführt, um eine Kompatibilität mit der Codierung des Schriftzeichensatzes der Deutschen Referenz-Version (DRV) des 7-Bit-Code zu ermöglichen. Ansonsten ist die DRV8 identisch mit der ARV8.

Bei der Deutschen Referenz-Version des 8-Bit-Code (Code-Tabelle 2) sind die Zeichen in den Spalten 02 bis 07 identisch mit den Zeichen in den Spalten 2 bis 7 der Deutschen Referenz-Version des 7-Bit-Code nach DIN 66 003.

Die Spalten 00 und 01 sowie 08 und 09 sind unbenutzt, solange sie nicht jeweils von einem genormten und mit entsprechenden ESC-Folgen definierten Steuerzeichensatz belegt werden.

Die Spalten 10 bis 15 sind mit Sonderzeichen sowie nationalen Buchstaben und Buchstaben mit diakritischen Zeichen oder Akzenten belegt, die in verschiedenen europäischen Sprachen verwendet werden, sowie mit jenen Schriftzeichen, die in den Spalten 02 bis 07 durch die deutschen nationalen Schriftzeichen verdrängt wurden.

Diese Version wird eindeutig definiert durch die ESC-Folgen

CO-Satz: ESC 02/01 F
G0-Satz: ESC 02/08 04/11
C1-Satz: ESC 02/02 F
G1-Satz: ESC .../.../.../...¹⁾

Dabei steht (F) jeweils für eine der Bitkombinationen von 03/00 bis 07/14. Nähere Einzelheiten sind in DIN 66 203 angegeben.

Die Normbezeichnung für die Deutsche Referenz-Version des 8-Bit-Code lautet:

Code DIN 66 303 – DRV8

5.3 Multibyte Version des 8-Bit-Code mit diakritischen Zeichen (MBV8)

Bei der Multibyte Version des 8-Bit-Code mit diakritischen Zeichen (Code-Tabelle 3) sind die Zeichen in den Spalten 02 bis 07 identisch mit den Zeichen in den Spalten 02 bis 07 der Internationalen Referenz-Version des 7-Bit-Code nach DIN 66 003.

Die Spalten 00 und 01 sowie 08 und 09 sind unbenutzt, solange sie nicht jeweils von einem genormten und mit entsprechenden ESC-Folgen definierten Steuerzeichensatz belegt werden.

Die Spalten 10 bis 15 sind mit Sonderzeichen, diakritischen Zeichen und Akzenten sowie mit nationalen Buchstaben belegt, die in verschiedenen europäischen Sprachen verwendet werden. Es werden nur solche Sprachen berücksichtigt, die sich den Schriftzeichen des lateinischen Alphabets bedienen. Die Spalten 10 bis 15 entsprechen dem Ergän-

¹⁾ Die ESC-Folge steht erst nach erfolgter Registrierung zur Verfügung. Nähere Auskunft erteilt der Normenausschuß Informationsverarbeitungssysteme (NI) im DIN, Burggrafenstraße 6, 1000 Berlin 30.

zungszeichensatz der Textkommunikation nach der Internationalen Norm ISO 6937/2-1983, wobei die Zeichen „nicht trennbares Leerzeichen“ (Platz 10/00), „Punkt untergesetzt“ (Platz 12/09), „unterbrochener senkrechter Strich“ (Platz 13/07), „Nicht-Zeichen“ (Platz 13/06) und „Silbentrennungsstrich“ (Platz 15/15) hinzugefügt wurden. Der Schriftzeichenvorrat der MBV8 ist größer als der Schriftzeichensatz, er ergibt sich aus den Tabellen 5 und 6. Bei Buchstaben mit diakritischen Zeichen oder Akzenten werden ihre Komponenten einzeln codiert, wobei das diakritische Zeichen oder der Akzent vor dem Buchstaben zu codieren ist. Die Akzente und diakritischen Zeichen, die mit Buchstaben kombiniert werden können, sind in der Spalte 12 codiert und bewirken keinen Vorschub der Aktuellen Zeichenstelle.

Diese Version wird eindeutig definiert durch die ESC-Folgen

C0-Satz: ESC 02/01 F

G0-Satz: ESC 02/08 04/00

C1-Satz: ESC 02/02 F

G1-Satz: ESC .../.../...¹⁾

Dabei steht (F) jeweils für eine der Bitkombinationen von 03/00 bis 07/14. Nähere Einzelheiten sind in DIN 66 203 angegeben.

Die Normbezeichnung für die Multibyte Version des 8-Bit-Code mit diakritischen Zeichen lautet:

Code DIN 66 303 – MBV8

¹⁾ Siehe Seite 2

Code-Tabelle 1. **Allgemeine Referenz-Version des 8-Bit-Code (ARV8)**

				b_8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	
				b_7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	0	0	0	00			SP	0	@	P	.	p			NBSP	°	À	Đ	à	đ
0	0	0	1	01			!	1	A	Q	a	q			i	±	Á	Ñ	á	ñ
0	0	1	0	02			"	2	B	R	b	r			¢	²	Â	Ò	â	ò
0	0	1	1	03			#	3	C	S	c	s			£	³	Ã	Ó	ã	ó
0	1	0	0	04			\$	4	D	T	d	t			¤	'	Ä	Ô	ä	ô
0	1	0	1	05			%	5	E	U	e	u			¥	µ	Å	Ö	å	ö
0	1	1	0	06			&	6	F	V	f	v			¦	¶	Æ	Ö	æ	ö
0	1	1	1	07			'	7	G	W	g	w			§	·	Ç	×	ç	÷
1	0	0	0	08			(	8	H	X	h	x			¨	,	È	Ø	è	ø
1	0	0	1	09			)	9	I	Y	i	y			©	¹	É	Ù	é	ù
1	0	1	0	10			*	:	J	Z	j	z			ª	º	Ê	Ú	ê	ú
1	0	1	1	11			+	;	K	[	k	{			«	»	Ë	Û	ë	û
1	1	0	0	12			,	<	L	\	l				¬	¼	Ì	Ü	ì	ü
1	1	0	1	13			-	=	M	]	m	}			SHY	½	Í	Ý	í	ý
1	1	1	0	14			.	>	N	^	n	~			®	¾	Î	Ï	î	ï
1	1	1	1	15			/	?	O	_	o				-	¿	Ï	ß	ï	ÿ

Für die Spalten 00 und 01 ist der zu verwendende C0-Satz durch die entsprechende ESC-Folge zu definieren.
Für die Spalten 08 und 09 ist der zu verwendende C1-Satz durch die entsprechende ESC-Folge zu definieren.

Code-Tabelle 2. Deutsche Referenz-Version des 8-Bit-Code (DRV8)

				b_8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	
				b_7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
				b_6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
				b_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
b_4	b_3	b_2	b_1		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	0	0	0	00			SP	0	§	P	`	p			NBSP	°	À	Ð	à	ð
0	0	0	1	01			!	1	A	Q	a	q			ı	±	Á	Ñ	á	ñ
0	0	1	0	02			"	2	B	R	b	r			¢	²	Â	Ò	â	ò
0	0	1	1	03			#	3	C	S	c	s			£	³	Ã	Ó	ã	ó
0	1	0	0	04			\$	4	D	T	d	t			¤	'	[	Ô	{	ô
0	1	0	1	05			%	5	E	U	e	u			¥	µ	Å	Ö	å	ö
0	1	1	0	06			&	6	F	V	f	v			¦	¶	Æ	\	æ	
0	1	1	1	07			'	7	G	W	g	w			@	·	Ç	×	ç	÷
1	0	0	0	08			(	8	H	X	h	x			¨	,	È	Ø	è	ø
1	0	0	1	09			)	9	I	Y	i	y			©	¹	É	Ù	é	ù
1	0	1	0	10			*	:	J	Z	j	z			ª	º	Ê	Ú	ê	ú
1	0	1	1	11			+	;	K	Ä	k	ä			«	»	Ë	Û	ë	ü
1	1	0	0	12			,	<	L	Ö	l	ö			¬	¼	Ï	]	ì	}
1	1	0	1	13			-	=	M	Ü	m	ü			SHY	½	Í	Ý	í	ý
1	1	1	0	14			.	>	N	^	n	ß			®	¾	Î	Þ	î	þ
1	1	1	1	15			/	?	O	_	o				-	¿	Ï	~	ï	ÿ

Für die Spalten 00 und 01 ist der zu verwendende C0-Satz durch die entsprechende ESC-Folge zu definieren.

Für die Spalten 08 und 09 ist der zu verwendende C1-Satz durch die entsprechende ESC-Folge zu definieren.

Code-Tabelle 3. Multibyte Version des 8-Bit-Code mit diakritischen Zeichen (MBV8)

				b_8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	
				b_7	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	
				b_6	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	
				b_5	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	
b_4	b_3	b_2	b_1		00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
0	0	0	0	00			SP	0	@	P	`	p			NBSP ²⁾	°		—	Ω	K
0	0	0	1	01			!	1	A	Q	a	q			i	±	·	¹	Æ	æ
0	0	1	0	02			"	2	B	R	b	r			¢	²	·	®	Ð	ð
0	0	1	1	03			#	3	C	S	c	s			£	³	·	©	ä	ö
0	1	0	0	04			¤	4	D	T	d	t			\$	×	·	™	⌘	⌚
0	1	0	1	05			%	5	E	U	e	u			¥	μ	·	♢		⌞
0	1	1	0	06			&	6	F	V	f	v				¶	·	↵ ²⁾	IJ	ij
0	1	1	1	07			'	7	G	W	g	w			§	·	·	¡ ²⁾	L	l
1	0	0	0	08			(	8	H	X	h	x				÷	·		Ł	ł
1	0	0	1	09			)	9	I	Y	i	y			·	·	·		Ø	ø
1	0	1	0	10			*	:	J	Z	j	z			"	"	°		Œ	œ
1	0	1	1	11			+	;	K	[	k	{			«	»	·		º	ß
1	1	0	0	12			,	<	L	\	l				←	¼	—	½	Þ	þ
1	1	0	1	13			-	=	M	]	m	}			↑	½	"	¾	ƒ	ƒ
1	1	1	0	14			·	>	N	^	n	~			→	¾	·	⅝	ŋ	ŋ
1	1	1	1	15			/	?	O	_	o				↓	¿	·	⅞	'n	²⁾ SHY

Für die Spalten 00 und 01 ist der zu verwendende C0-Satz durch die entsprechende ESC-Folge zu definieren.

Für die Spalten 08 und 09 ist der zu verwendende C1-Satz durch die entsprechende ESC-Folge zu definieren.

Die in den Spalten 10 bis 15 dunkel getönten (gepunkteten) Plätze sind einer künftigen Normung vorbehalten und dürfen nicht benutzt werden.

²⁾ Die Belegung dieser Plätze ist international noch nicht abgesichert.

5.4 Erklärung der in den Code-Tabellen benutzten internationalen Kurzzeichen

Siehe Tabelle 4. Die in Klammern in kursiver Schrift hinzugefügten englischen Benennungen sind den Internationalen Normen und Entwürfe (ISO 6937/2 – 1983 und ISO/DIS 8859/1.2 – 1986) entnommen.

Tabelle 4. Zeichen der Code-Tabellen 1 bis 3

Platz (Spalte/Zeile)	Kurzzeichen	Benennung
02/00	SP	Leerzeichen (<i>Space</i>)
10/00	NBSP	nicht trennbares Leerzeichen
in Code-Tabelle 1 und 2: 10/13 3: 15/15	SHY	Silbentrennungsstrich

5.5 Erklärung der in den Code-Tabellen benutzten internationalen Kurzzeichen und Schriftzeichen

Siehe Tabellen 5 und 6. Die Schriftzeichen in der Tabelle 5 sind in der Reihenfolge der Plätze in der Code-Tabelle 1 aufgeführt und denen der Code-Tabellen 2 und 3 gegenüber gestellt. Schriftzeichen, die in den Code-Tabellen 1 und 2 nicht enthalten sind, wurden in der Reihenfolge der Plätze in der Code-Tabelle 3 aufgeführt. Die Tabelle 6 führt die Kombinationen von lateinischen Grundbuchstaben mit diakritischen Zeichen oder Akzenten aus der Spalte 12 der Code-Tabelle 3 auf. Die in den Tabellen 5 und 6 benutzte Kennung für die Schriftzeichen ist im Anhang A beschrieben.

Tabelle 5. Schriftzeichen der Code-Tabellen 1 bis 3

Kennung	Schrift- zeichen (Kurzbe- zeichnung)	Platz (Spalte/Zeile) in Code-Tabelle			Benennung
		1	2	3	
SP01	SP	02/00	02/00	02/00	Leerzeichen
SP02	!	02/01	02/01	02/01	Ausrufezeichen
SP04	"	02/02	02/02	02/02	Anführungszeichen
SM01	#	02/03	02/03	02/03	Nummernzeichen
SC03	\$	02/04	02/04	10/04	Dollar-Zeichen
SM02	%	02/05	02/05	02/05	Prozentzeichen
SM03	&	02/06	02/06	02/06	kaufmännisches Und
SP05	'	02/07	02/07	02/07	Apostroph (Auslassungszeichen)
SP06	(	02/08	02/08	02/08	runde Klammer links
SP07	)	02/09	02/09	02/09	runde Klammer rechts
SM04	*	02/10	02/10	02/10	Stern
SA01	+	02/11	02/11	02/11	Pluszeichen
SP08	,	02/12	02/12	02/12	Komma (Beistrich)
SP10	-	02/13	02/13	02/13	Mittestrich (Silbentrennungsstrich, Bindestrich, Gedankenstrich) Minuszeichen
SP11	.	02/14	02/14	02/14	Punkt
SP12	/	02/15	02/15	02/15	Schrägstrich
ND10	0	03/00	03/00	03/00	Ziffer Null
ND01	1	03/01	03/01	03/01	Ziffer Eins
ND02	2	03/02	03/02	03/02	Ziffer Zwei
ND03	3	03/03	03/03	03/03	Ziffer Drei
ND04	4	03/04	03/04	03/04	Ziffer Vier
ND05	5	03/05	03/05	03/05	Ziffer Fünf