

GSM Alphabet

Contents

1. Overview
2. GSM default alphabet
3. GSM extended character set

1. Overview

All text contained in an SMS message either uses the GSM character set, UCS2 or national language shift tables (see [GSM 03.38](#) for details). Up to 160 x GSM characters or 70 x UCS2 characters will fit into a single SMS. Where a text message is larger than the capacity of a single SMS then segmentation and reassembly headers in each SMS are used to allow the message to span multiple SMS.

2. GSM default alphabet

The GSM default alphabet is shown in the following table.

	0-	1-	2-	3-	4-	5-	6-	7-
-0	@		SP	0	i	P	ç	p
-1	£	_	!	1	A	Q	a	q
-2	\$		“	2	B	R	b	r
-3	¥		#	3	C	S	c	s
-4	è		œ	4	D	T	d	t
-5	é		%	5	E	U	e	u
-6	ù		&	6	F	V	f	v
-7	ì		‘	7	G	W	g	w
-8	ò		(8	H	X	h	x

-9	Ç)	9	l	Y	i	y
-A	LF		*	:	J	Z	j	z
-B	Ø		+	;	K	Ä	k	ä
-C	ø	Æ	,	<	L	Ö	l	ö
-D	CR	æ	-	=	M	Ñ	m	ñ
-E	Å	ß	.	>	N	Ü	n	ü
-F	å	É	/	?	O	§	o	à

3. GSM extended character set

Further characters are supported using an extended character set that is accessed using the escape character (0x1B). These characters are as follows:

	0-	1-	2-	3-	4-	5-	6-	7-
-0								
-1								
-2								
-3								

-4		^						
-5						€		
-6								
-7								
-8			{					
-9			}					
-A								
-B								
-C				[
-D				~				
-E]				
-F			\					

Sources:

- ETSI TS 100 900 (GSM 03.38 § 6.2.1) / 3GPP TS 23.038
- https://en.wikipedia.org/wiki/GSM_03.38

Notes:

- Based on the above tables showing the GSM character set, to represent the Euro (€) character the values 0x1B 0x65 are used.
- In GSM mobile networks up to 160 x GSM characters (i.e. $160 \times 7\text{-bits} = 1120$ bits) or 70 x UCS2 characters (i.e. $70 \times 16\text{-bits} = 1120$ bits) will fit into a single SMS.