

3.6 Code System

3.6.1 Character encoding system

A character encoding system herein refers to a system to encode a set of characters used in NACCS EDI messages. Note that this specific character encoding system is applicable only to NACCS EDI messages. This character encoding system, therefore, is not applicable to communication protocol headers nor trailers which are added upon transmission or receipt.

In addition, "4.6.2 Attachment file" shows the character codes and filename extension used in attachment file names.

3.6.1.1 Encoded characters used in outbound messages (processing request messages)

(1) One-byte characters (attribute "n" and "an")

Table 3.6.1 below shows a one-byte character set in the bold frame used in creating processing request messages to be sent by users (attributes "n" and "an").

Table 3.6.1 One-byte Characters Allowed to be Used in Processing Request Messages

Bits								Column							
b7	b6	b5	B4	b3	b2	b1	row	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0			SP	0	@	P		
0	0	0	1	1			1			!	1	A	Q		
0	0	1	0	2			2			"	2	B	R		
0	0	1	1	3			3			#	3	C	S		
0	1	0	0	4			4			\$	4	D	T		
0	1	0	1	5			5			%	5	E	U		
0	1	1	0	6			6			&	6	F	V		
0	1	1	1	7			7			'	7	G	W		
1	0	0	0	8			8			(8	H	X		
1	0	0	1	9			9)	9	I	Y		
1	0	1	0	A			A	LF		*	:	J	Z		
1	0	1	1	B			B			+	;	K	[
1	1	0	0	C			C			,	<	L	¥		
1	1	0	1	D			D	CR		-	=	M]		
1	1	1	0	E			E			.	>	N	^		
1	1	1	1	F			F			/	?	O	_		

(Note) SP denotes a space.

(Note 1) Invalid characters: Half-width katakana characters, "\$", "[", "]", "^", and "_".

(Note 2) "¥" can be used only in "Invoice and Packing List Information Registration (IVA)".

(2) One-byte characters (attribute "sn")

One-byte characters (attribute "sn") that users can use to create processing request messages should be the range indicated within the bold frame on Table 3.6.2.

Table 3.6.2 One-byte Characters Allowed to be Used in Processing Request Messages (Exception)

								0	0	0	0	1	1	1	1	
								0	0	1	1	0	0	1	1	
								0	1	0	1	0	1	0	1	
								Column								
Bits	b7	B6	b5	B4	b3	b2	b1	Row	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0			SP	0	@	P	`	p
0	0	0	1	1	1	1	1	1			!	1	A	Q	a	q
0	0	1	0	2	2	2	2	2			"	2	B	R	b	r
0	0	1	1	3	3	3	3	3			#	3	C	S	c	s
0	1	0	0	4	4	4	4	4			\$	4	D	T	d	t
0	1	0	1	5	5	5	5	5			%	5	E	U	e	u
0	1	1	0	6	6	6	6	6			&	6	F	V	f	v
0	1	1	1	7	7	7	7	7			'	7	G	W	g	w
1	0	0	0	8	8	8	8	8			(8	H	X	h	x
1	0	0	1	9	9	9	9	9)	9	I	Y	i	y
1	0	1	0	A	A	A	A	A	LF		*	:	J	Z	j	z
1	0	1	1	B	B	B	B	B			+	;	K	[k	{
1	1	0	0	C	C	C	C	C			,	<	L	¥	l	
1	1	0	1	D	D	D	D	D	CR		-	=	M]	m	}
1	1	1	0	E	E	E	E	E			.	>	N	^	n	~
1	1	1	1	F	F	F	F	F			/	?	O	_	o	

Refer to "4.6.2 Attachment file" for attachment file names available for NACCS.

(3) Two-byte characters

Characters that users can use to create processing request messages with two-byte characters shall comply with JIS X 0208:1997. Users can only use JIS Level 1 and Level 2 Japanese kanji characters. The EUC-JP encoding shall be used.

(4) Mixture of two-byte and one-byte characters in processing request messages

In processing request messages, two-byte characters are allowed only in particular elements (attribute: j). Two-byte characters can be mixed with one-byte characters in elements where the use of two-byte characters is allowed. One-byte characters that are used in combination with two-byte characters are those in the bold frame given in Table 3.6.3.

Character set which can be used in EDIFACT messages should be the Level A Character Set as well as the 2 characters "#" and "@"

Table 3.2.1 described in 3.2.1.1.5 shows Level A Character Set.

Table 3.6.3 One-byte Characters Allowed to be Used with Two-byte Characters

Bits	b7	b6	b5	b4	b3	b2	b1	row	Column							
									0	1	2	3	4	5	6	7
									0	0	0	0	1	1	1	1
									0	0	1	1	0	0	1	1
									0	1	0	1	0	1	0	1
									0	1	2	3	4	5	6	7
	0	0	0	0	0	0	0	0			SP	0	@	P	`	p
	0	0	0	1	1			1			!	1	A	Q	a	q
	0	0	1	0				2			"	2	B	R	b	r
	0	0	1	1				3			#	3	C	S	c	s
	0	1	0	0				4			\$	4	D	T	d	t
	0	1	0	1				5			%	5	E	U	e	u
	0	1	1	0				6			&	6	F	V	f	v
	0	1	1	1				7			'	7	G	W	g	w
	1	0	0	0				8			(8	H	X	h	x
	1	0	0	1				9)	9	I	Y	i	y
	1	0	1	0			A	A	LF		*	:	J	Z	j	z
	1	0	1	1			B	B			+	;	K	[k	{
	1	1	0	0			C	C			,	<	L	¥	l	
	1	1	0	1			D	D	CR		-	=	M]	m	}
	1	1	1	0			E	E			.	>	N	^	n	~
	1	1	1	1			F	F			/	?	O	_	o	

3.6.1.2 Encoded characters used in inbound messages (processing result messages)

(1) One-byte characters (attribute "n" and "an")

One-byte characters (attribute "n" and "an") that are used in processing result messages (processing result notification messages, output information messages) transmitted from the NACCS Center server to users shall be those in the bold frame given in Table 3.6.4.

Character set which can be used in EDIFACT messages should be the Level A Character Set as well as the 2 characters "#" and "@" ("#" and "@" are not treated as "errors".)

Table 3.2.1 described in 3.2.1.1.5 shows Level A Character Set.

Table 3.6.4 One-byte Characters Allowed to be Used in Processing Result Messages

Bits	b7	b6	b5	b4	b3	b2	b1	row	Column									
									0	1	2	3	4	5	6	7		
									0	0	0	0	1	1	1	1		
									0	0	1	1	0	0	1	1		
									0	1	0	1	0	1	0	1		
									0	1	0	0	4	5	6	7		
								0	0	0	0	0	SP	@	P			
								0	0	0	1	1	!	A	Q			
								0	0	1	0	2	"	B	R			
								0	0	1	1	3	#	C	S			
								0	1	0	0	4	\$	D	T			
								0	1	0	1	5	%	E	U			
								0	1	1	0	6	&	F	V			
								0	1	1	1	7	'	G	W			
								1	0	0	0	8	(H	X			
								1	0	0	1	9)	I	Y			
								1	0	1	0	A	LF	*	:	J	Z	
								1	0	1	1	B		+	;	K	[
								1	1	0	0	C		,	<	L	¥	
								1	1	0	1	D	CR	-	=	M]	
								1	1	1	0	E		.	>	N	^	
								1	1	1	1	F		/	?	O	_	

(Note) SP denotes a space.

(2) One-byte characters (attribute "sn")

One-byte characters (attribute "sn") that are used in processing result messages (processing result notification messages, output information messages) transmitted from the NACCS Center server to users shall be those in the bold frame given in Table 3.6.5.

Table 3.6.5 One-byte Characters Allowed to be Used in Processing Result Messages (Exception)

Bits	b7	B6	b5	B4	b3	b2	b1	row	Column								
									0	1	2	3	4	5	6	7	
									0	0	0	0	1	1	1	1	
									0	0	1	1	0	0	1	1	
									0	1	0	1	0	1	0	1	
									0	1	2	3	4	5	6	7	
	0	0	0	0	0	0	0	0			SP	0	@	P	`	p	
	0	0	0	1	1			1			!	1	A	Q	a	q	
	0	0	1	0	2			2			"	2	B	R	b	r	
	0	0	1	1	3			3			#	3	C	S	c	s	
	0	1	0	0	4			4			\$	4	D	T	d	t	
	0	1	0	1	5			5			%	5	E	U	e	u	
	0	1	1	0	6			6			&	6	F	V	f	v	
	0	1	1	1	7			7			'	7	G	W	g	w	
	1	0	0	0	8			8			(8	H	X	h	x	
	1	0	0	1	9			9)	9	I	Y	i	y	
	1	0	1	0	A			A	LF		*	:	J	Z	j	z	
	1	0	1	1	B			B			+	;	K	[k	{	
	1	1	0	0	C			C			,	<	L	¥	l		
	1	1	0	1	D			D	CR		-	=	M]	m	}	
	1	1	1	0	E			E			.	>	N	^	n	~	
	1	1	1	1	F			F			/	?	O	_	o		

(3) Two-byte characters

Characters that users can use to create processing request messages with two-byte characters shall comply with JIS X 0208:1997. Users can only use JIS Level 1 and Level 2 Japanese kanji characters. The EUC-JP encoding shall be used.

(4) Mixture of two-byte and one-byte characters in processing result messages

Two-byte characters are used in some types of output messages because those contain elements requiring Japanese characters and symbols (attribute: j). In such cases, the EUC-JP encoding shall be used as the Character code. However, a mixture of one-byte characters is allowable. One-byte characters that are used in combination with two-byte characters are those in the bold frame given in Table 3-6-6. In addition, the packaged software provided by NACCS automatically converts EUC-JP character data to Shift-JIS character data, EUC-JP characters, therefore, are displayed as Shift-JIS characters on computers.

Character set which can be used in EDIFACT messages should be the Level A Character Set as well as the 2 characters "#" and "@" ("#" and "@" are not treated as "errors".)

Table 3.2.1 described in 3.2.1.1.5 shows Level A Character Set.

Table 3-6-6 One-byte character set in output elements with Japanese notation

Bits	b7	b6	b5	b4	b3	b2	b1	row	Column							
									0	1	2	3	4	5	6	7
									0	0	0	0	1	1	1	1
									0	0	1	1	0	0	1	1
									0	1	0	1	0	1	0	1
									0	1	2	3	4	5	6	7
				0	0	0	0	0			SP	0	@	P	`	p
				0	0	0	1	1			!	1	A	Q	a	q
				0	0	1	0	2			"	2	B	R	b	r
				0	0	1	1	3			#	3	C	S	c	s
				0	1	0	0	4			\$	4	D	T	d	t
				0	1	0	1	5			%	5	E	U	e	u
				0	1	1	0	6			&	6	F	V	f	v
				0	1	1	1	7			'	7	G	W	g	w
				1	0	0	0	8			(8	H	X	h	x
				1	0	0	1	9)	9	I	Y	i	y
				1	0	1	0	A		LF	*	:	J	Z	j	z
				1	0	1	1	B			+	;	K	[k	{
				1	1	0	0	C			,	<	L	¥	l	
				1	1	0	1	D		CR	-	=	M]	m	}
				1	1	1	0	E			.	>	N	^	n	~
				1	1	1	1	F			/	?	O	_	o	

3.6.2 Garbled characters

In NACCS, some procedures involve descriptions in two-byte characters (Kanji, etc.). Due to a mixture of two-byte and one-byte characters allowed in such elements and the absence of byte sensitivity, platform dependent characters and symbols may be garbled and illegible, depending on the type of gateway computer.

In addition, if the user's gateway computer adds shift codes, the message length may be altered.