

Thai Input and Output Methods

Translated from "Computer and Thai Language"

Translated by Chamrat Atsawaprecha

February 12, 1992



Introduction

This document discusses about Thai character classification which resulted from requirements for data processing. It also defines the behaviour of system programs and application programs when working with Thai characters. Futhermore, this document defined the common behavior for input and output methods of Thai and English.

The common behaviour is defined such that input method and output method of any application programs running on a Thai system will behave the same for any given Thai or English character strings.

1 Character Sets

1.1 Basic Character Set

The character in this character set is 8-bit coded. The character codes have the values 0-255 (UCHAR or unsigned char). Figure 1-1 shows characters in the Basic Character Set which shall be called TACTIS (Thai API Consortium/Thai Industrial Standard) code. The TACTIS code is the result from combining ISO 646-1983 (7-bit coded character set for information interchange) with TIS 620-2533 (Standard for Thai character code for computer). The TACTIS code is compatible with specifications defined in ISO-2022 (Code extension techniques). Therefore, TACTIS code can be used for international data transmission properly.

TACTIS code has been well known in Thailand computer industries because the combining of ISO 646-1983 and TIS 620-2533 has been done correctly. Therefore, when transmitting data accross network, characters that are reserved for future use should not be used.

Character Set mentioned in this document unspecifically shall refer to the Basic Character Set (TACTIS code).

1.2 Auxiliary Character Set

There should be at least one auxiliary character set but there can be more than one. The rules are that the auxiliary character set must use character codes 32-126 and 161-254 only, and each character uses one display cell.

In a MS-DOSTM sytem, the auxiliary character set that it must have is Code Page 437.

Basic Character Set

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	DLE	SP	0	@	P	`	p				ฐ	ภ	ะ	เ	๐
1	SOH	DC1	!	1	A	Q	a	q			ก	ท	ม	๑	แ	๑
2	STX	DC2	"	2	B	R	b	r			ข	ฒ	ย	า	โ	๒
3	ETX	DC3	#	3	C	S	c	s			ช	ณ	ร	ำ	ใ	๓
4	EOT	DC4	\$	4	D	T	d	t			ค	ต	ฤ	๒	ไ	๔
5	ENQ	NAK	%	5	E	U	e	u			ค	ต	ล	๓	า	๕
6	ACK	SYN	&	6	F	V	f	v			ฆ	ถ	ภ	๓	า	๖
7	BEL	ETB	'	7	G	W	g	w			ง	ท	ว	๓	๓	๗
8	BS	CAN	(8	H	X	h	x			จ	ฐ	ศ	๑	๑	๘
9	HT	EM)	9	I	Y	i	y			ฉ	น	ษ	๑	๑	๙
A	LF	SUB	*	:	J	Z	j	z			ช	บ	ส	๑	๑	๑๑
B	VT	ESC	+	;	K	[k	{			ช	ป	ห	๑	๑	๑๑๑
C	FF	FS	,	<	L	\	l				ณ	ผ	พ	word sep.	๑	
D	CR	GS	-	=	M]	m	}			ญ	ฝ	อ	๑	๑	
E	SO	RS	.	>	N	^	n	~			ญ	พ	ฮ	๑	๑	
F	SI	US	/	?	O	_	o	DEL			ญ	ฟ	า	฿	๑	

ISO 646-1983
 TIS 620-2533
 Reserved for future use

word sep. = word separator character

The circle is not part of the glyph; it is meant to help indicate above/below level of vowels/tonemarks/diacritics.

Figure 1 -1 Basic Character Set (TACTIS code)

2 Character Classification

Thai character classification in this standard is meant to ease computer processing when dealing with displaying and input-sequence checking Thai characters. This classification, by all mean, is not related to Thai linguistic. However, the classification will help displaying Thai characters correctly according to the Thai language.

2.1 Control characters (CTRL). These are nondisplayable characters being used as control codes for output display or data communication. Sixty-six control characters are (0x00) to (0x1F), (0x7F), (0x80) to (0x9F), (0xFF).

2.2 Consonants (CONS). Forty-four consonants are (0xA1)-(0xCE).

ก ข ฃ ค ฅ ฉ ง จ ฉ ช ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท
ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ฤ ล ฦ ว ศ ษ ส ห พ อ ฮ

2.3 Vowels (-V). Vowels are divided into four groups.

- **Leading Vowels** (LV). These vowels are written leading the consonants. Five leading vowels are ๑ (0xE0), ๒ (0xE1), ๓ (0xE2), ๔ (0xE3), and ๕ (0xE4).
- **Following Vowels** (FV). These vowels are written following the consonants. Six following vowels are divided into four normal following vowels and two special following vowels.
Four normal following vowels are ๖ (0xD0), ๗ (0xD2), ๘ (0xD3) and ๙ (0xE5).
Two special following vowels are ๑๐ (0xC4) and ๑๑ (0xC6).
- **Below Vowels** (BV). These vowels are written below the consonants. Two below vowels are ๑๒ (0xD8) and ๑๓ (0xD9).
- **Above Vowels** (AV). These vowels are written above the consonants. Five above vowels are ๑๔ (0xD4), ๑๕ (0xD5), ๑๖ (0xD6), ๑๗ (0xD7) and ๑๘ (0xD1).

2.4 Tonemarks (TONE). Four tonemarks are ' (0xE8), ¨ (0xE9), ˇ (0xEA) and + (0xEB).

2.5 Diacritics (-D). Diacritics are divided into two groups.

- *Above Diacritics (AD).* These diacritics are written above initial or final consonants. Four above diacritics are ° (0xED), ¨ (0xEC), ˇ (0xE7) and ¨ (0xEE).
- *Below Diacritic (BD).* Below diacritic is written below final or clustered consonants. The below diacritic is . (0xDA).

2.6 Non-composables (NON). These are characters that do not fit into five classes above. These characters are not composable. Above vowels, below vowels, tonemarks, above diacritics and below diacritics cannot compose with these group of 119 characters. Non-composables are divided into seven groups.

- *Graphic Characters.* According to ISO 646-1983, there are 94 graphic characters. These graphic characters are subdivided into 52 English alphabetic characters (A B C D E F G H I J K L M N O P Q R S T U V W X Y Z a b c d e f g h i j k l m n o p q r s t u v w x y z), 10 digits (0 1 2 3 4 5 6 7 8 9), and 32 special characters which include ! " # \$ % & ' () * + , - . / : ; < = > ? @ [\] ^ _ ` { | } ~
- *Space .* Character code is (0x20).
- *No-break Space.* Character code is (0xA0).
- *Thai Digits.* According to TIS 620-2533, ten Thai digits are ๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙ .
- *Thai Special Characters.* According to TIS 620-2533, six Thai special characters are ๑ ๒ ๓ ๔ ๕ ๖

- *Word Break*. Character code is (0xDC). Word break character is a non-printing character. It is to be used in breaking words in Thai sentences. Thus, application like Thai spell checkers can make use of word break characters to simplify processing.
- *Reserved Characters*. Six of them are (0xDB), (0xDD), (0xDE), (0xFC), (0xFD) and (0xFE).

In order to describe Thai input/output methods, character classification for some classes (FV, BV, AV, and AD) have been re-classified into subclass such as FV1, FV2 and FV3. See Table 2-1 for detail.

<i>Class</i>	<i>Numbers</i>	<i>Description</i>																
CTRL	66	ISO 646-1983 control codes (0x00)-(0x1F), (0x7F), (0x80-0x9F), (0xFF)																
NON	119	<p>According to ISO 646-1983, character codes (0x20-0x7E, 0xA0).</p> <p>According to TIS 620-2533, the following characters.</p> <table style="margin-left: 40px;"> <tr> <td>๑</td> <td>(0xCF)</td> </tr> <tr> <td>Thai word break</td> <td>(0xDC)</td> </tr> <tr> <td>฿</td> <td>(0xDF)</td> </tr> <tr> <td>๑</td> <td>(0xE6)</td> </tr> <tr> <td>●</td> <td>(0xEF)</td> </tr> <tr> <td>๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙</td> <td>(0xF0)-(0xF9)</td> </tr> <tr> <td>๑</td> <td>(0xFA)</td> </tr> <tr> <td>๑</td> <td>(0xFB)</td> </tr> </table> <p>Character codes not being used by any standard (0xDB), (0xDD), (0xDE), (0xFC), (0xFD), (0xFE).</p>	๑	(0xCF)	Thai word break	(0xDC)	฿	(0xDF)	๑	(0xE6)	●	(0xEF)	๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙	(0xF0)-(0xF9)	๑	(0xFA)	๑	(0xFB)
๑	(0xCF)																	
Thai word break	(0xDC)																	
฿	(0xDF)																	
๑	(0xE6)																	
●	(0xEF)																	
๐ ๑ ๒ ๓ ๔ ๕ ๖ ๗ ๘ ๙	(0xF0)-(0xF9)																	
๑	(0xFA)																	
๑	(0xFB)																	
CONS	44	<p>ก ข ฃ ค ฅ ง จ ฉ ช ซ ฌ ญ ฎ ฏ ฐ ฑ ฒ ณ ด ต ถ ท ธ น บ ป ผ ฝ พ ฟ ภ ม ย ร ล ฤ ฦ ว ศ ษ ส ห พ อ ฮ</p>																

<i>Class</i>	<i>Numbers</i>	<i>Description</i>
LV	5	๓ (0xE0), ๔ (0xE1), ๕ (0xE2), ๖ (0xE3), ๗ (0xE4)
FV1	3	๘ (0xD0), ๙ (0xD2), ๐ (0xD3)
FV2	1	๑ (0xE5)
FV3	2	๒ (0xC4), ๓ (0xC6)
BV1	1	๔ (0xD8)
BV2	1	๕ (0xD9)
BD	1	๖ (0xDA)
TONE	4	๗ (0xE8), ๘ (0xE9), ๙ (0xEA), ๐ (0xEB)
AD1	2	๑ (0xED), ๒ (0xEC)
AD2	1	๓ (0xE7)
AD3	1	๔ (0xEE)
AV1	1	๕ (0xD4)
AV2	2	๖ (0xD1), ๗ (0xD6)
AV3	2	๘ (0xD5), ๙ (0xD7)

Table 2 -1 Thai Character Classification

3 Properties of Characters in the Basic Character Set

Thai/English characters when being output in any computer systems must have common properties. The following section discusses the properties of characters and the API functions that check for these properties.

3.1 Character Type (`ctype`). There are sixteen types of characters in the basic character set. See table 3-1 for the values corresponding to each class of characters. The following is the API for function `TACctype`.

```
int TACctype (unsigned char ch);
```

This function returns `ctype` which has the value between 0-16 as in table ?.

<i>Class</i>	<i>ctype</i>
CTRL	0
NON	1
CONS	2
LV	3
FV1	4
FV2	5
FV3	6
BV1	7
BV2	8
BD	9
TONE	10
AD1	11
AD2	12
AD3	13
AV1	14
AV2	15
AV3	16

Table 3 -1 `ctype` value of each class of character

3.2 Character Level (`chlevel`). There are five character levels. Four levels are displayable and one level is non-displayable. The five levels are explained as follows.

3.3 Other Characters' Properties. Each character has its own properties which are summarized in table 3-2.

- Dead characters. Dead characters are characters in CTRL, BV, AV, BD, AD and TONE classes. Since a dead character (especially vowel, diacritic or tonemark) is usually composed with a consonant, after a dead character is typed, the cursor will not be advanced to the next character display cell.
- Consonants. Consonants are all characters in CONS class.
- Vowels. Vowels are all characters in LV, FV, BV, and AV classes.
- Tonemarks. Tonemarks are all characters in TONE class.
- Digits. Twenty digits are 0 1 2 3 4 5 6 7 8 9 and ଠ ଡ ଢ ଣ ଣ ଝ ଞ ଟ ଠ ଡ ଢ ଣ.

<i>class</i>	<i>chtype</i>	<i>chlevel</i>	<i>isdead</i>	<i>iscons</i>	<i>isvowel</i>	<i>istone</i>	<i>isdigit</i>	<i>ispunct</i>
CTRL	0	NONDISP	dead					
NON	1	BASE					0-9, ଠ-୯	some
CONS	2	BASE		cons				
LV	3	BASE			vowel			
FV1	4	BASE			vowel			
FV2	5	BASE			vowel			
FV3	6	BASE			vowel			
BV1	7	BELOW	dead		vowel			
BV2	8	BELOW	dead		vowel			
BD	9	BELOW	dead					
TONE	10	ABOVE	dead			tone		
AD1	11	ABOVE	dead					
AD2	12	ABOVE	dead					
AD3	13	ABOVE	dead					
AV1	14	ABOVE	dead		vowel			
AV2	15	ABOVE	dead		vowel			
AV3	16	ABOVE	dead		vowel			

Table 3 -2 Summary of basic properties of Basic Character Set and name of functions to check properties. Note that name of functions must begin with TAC.

3.4 Character classification table.

<i>Dec</i>	<i>Hex</i>	<i>Pos</i>	<i>Char</i>	<i>Name</i>	<i>Class</i>	<i>Level</i>	<i>Properties</i>
0	00	00/00	NUL	NULL	CTRL	NONDISP	
1	01	00/01	SOH	START OF HEADING	CTRL	NONDISP	
2	02	00/02	STX	START OF TEXT	CTRL	NONDISP	
3	03	00/03	ETX	END OF TEXT	CTRL	NONDISP	
4	04	00/04	EOT	END OF TRANSMISSION	CTRL	NONDISP	
5	05	00/05	ENQ	ENQUIRY	CTRL	NONDISP	
6	06	00/06	ACK	ACKNOWLEDGE	CTRL	NONDISP	
7	07	00/07	BEL	BELL	CTRL	NONDISP	
8	08	00/08	BS	BACKSPACE	CTRL	NONDISP	
9	09	00/09	HT	HORIZONTAL TABULATION	CTRL	NONDISP	
10	0A	00/10	LF	LINE FEED	CTRL	NONDISP	
11	0B	00/11	VT	VERTICAL TAB	CTRL	NONDISP	
12	0C	00/12	FF	FORM FEED	CTRL	NONDISP	
13	0D	00/13	CR	CARRIAGE RETURN	CTRL	NONDISP	
14	0E	00/14	SO	SHIFT OUT	CTRL	NONDISP	
15	0F	00/15	SI	SHIFT IN	CTRL	NONDISP	
16	10	01/00	DLE	DATA LINK ESCAPE	CTRL	NONDISP	
17	11	01/01	DC1	DEVICE CONTROL 1, XON	CTRL	NONDISP	
18	12	01/02	DC2	DEVICE CONTROL 2	CTRL	NONDISP	
19	13	01/03	DC3	DEVICE CONTROL 3, XOFF	CTRL	NONDISP	
20	14	01/04	DC4	DEVICE CONTROL 4	CTRL	NONDISP	
21	15	01/05	NAK	NEGATIVE ACKNOWLEDGE	CTRL	NONDISP	
22	16	01/06	SYN	SYNCHRONOUS IDLE	CTRL	NONDISP	
23	17	01/07	ETB	END OF TRANSMISSION BLOCK	CTRL	NONDISP	
24	18	01/08	CAN	CANCEL	CTRL	NONDISP	
25	19	01/09	EM	END OF MEDIUM	CTRL	NONDISP	
26	1A	01/10	SUB	SUBSTITUTE CHARACTER	CTRL	NONDISP	
27	1B	01/11	ESC	ESCAPE	CTRL	NONDISP	
28	1C	01/12	FS	INFORMATION SEPARATOR 4, (FILE SEPARATOR)	CTRL	NONDISP	
29	1D	01/13	GS	INFORMATION SEPARATOR 3, (GROUP SEPARATOR)	CTRL	NONDISP	
30	1E	01/14	RS	INFORMATION SEPARATOR 2, (RECORD SEPARATOR)	CTRL	NONDISP	
31	1F	01/15	US	INFORMATION SEPARATOR 1, (UNIT SEPARATOR)	CTRL	NONDISP	
32	20	02/00	SP	SPACE	NON	BASE	
33	21	02/01	!	EXCLAMATION MARK	NON	BASE	punct
34	22	02/02	"	QUOTATION MARK	NON	BASE	punct
35	23	02/03	#	NUMBER SIGN	NON	BASE	punct
36	24	02/04	\$	DOLLAR SIGN	NON	BASE	punct
37	25	02/05	%	PERCENT SIGN	NON	BASE	punct
38	26	02/06	&	AMPERSAND	NON	BASE	punct

<i>Dec</i>	<i>Hex</i>	<i>Pos</i>	<i>Char</i>	<i>Name</i>	<i>Class</i>	<i>Level</i>	<i>Properties</i>
39	27	02/07	'	APOSTROPHE	NON	BASE	punct
40	28	02/08	(LEFT PARANTHESIS	NON	BASE	punct
41	29	02/09)	RIGHT PARENTHESIS	NON	BASE	punct
42	2A	02/10	*	ASTERISK	NON	BASE	punct
43	2B	02/11	+	PLUS SIGN	NON	BASE	punct
44	2C	02/12	,	COMMA	NON	BASE	punct
45	2D	02/13	-	HYPHEN, MINUS SIGN	NON	BASE	punct
46	2E	02/14	.	FULL STOP	NON	BASE	punct
47	2F	02/15	/	SOLIDUS	NON	BASE	punct
48	30	03/00	0	DIGIT ZERO	NON	BASE	digit
49	31	03/01	1	DIGIT ONE	NON	BASE	digit
50	32	03/02	2	DIGIT TWO	NON	BASE	digit
51	33	03/03	3	DIGIT THREE	NON	BASE	digit
52	34	03/04	4	DIGIT FOUR	NON	BASE	digit
53	35	03/05	5	DIGIT FIVE	NON	BASE	digit
54	36	03/06	6	DIGIT SIX	NON	BASE	digit
55	37	03/07	7	DIGIT SEVEN	NON	BASE	digit
56	38	03/08	8	DIGIT EIGHT	NON	BASE	digit
57	39	03/09	9	DIGIT NINE	NON	BASE	digit
58	3A	03/10	:	COLON	NON	BASE	punct
59	3B	03/11	;	SEMICOLON	NON	BASE	punct
60	3C	03/12	<	LESS-THAN SIGN	NON	BASE	punct
61	3D	03/13	=	EQUAL SIGN	NON	BASE	punct
62	3E	03/14	>	GREATER-THAN SIGN	NON	BASE	punct
63	3F	03/15	?	QUESTION MARK	NON	BASE	punct
64	40	04/00	@	COMMERCIAL AT	NON	BASE	punct
65	41	04/01	A	CAPITAL LETTER A	NON	BASE	xdigit alpha
66	42	04/02	B	CAPITAL LETTER B	NON	BASE	xdigit alpha
.....
90	5A	05/10	Z	CAPITAL LETTER Z	NON	BASE	alpha
91	5B	05/11	[LEFT SQUARE BRACKET	NON	BASE	punct
92	5C	05/12	\	REVERSE SOLIDUS	NON	BASE	punct
93	5D	05/13]	RIGHT SQUARE BRACKET	NON	BASE	punct
94	5E	05/14	^	CIRCUMFLEX ACCENT	NON	BASE	punct
95	5F	05/15	_	LOWLINE, UNDERLINE	NON	BASE	punct
96	60	06/00	`	GRAVE ACCENT	NON	BASE	punct
97	61	06/01	a	SMALL LETTER a	NON	BASE	xdigit alpha
98	62	06/02	b	SMALL LETTER b	NON	BASE	xdigit alpha
.....
122	7A	07/10	z	SMALL LETTER z	NON	BASE	alpha

<i>Dec</i>	<i>Hex</i>	<i>Pos</i>	<i>Char</i>	<i>Name</i>	<i>Class</i>	<i>Level</i>	<i>Properties</i>
123	7B	07/11	{	LEFT CURLY BRACKET	NON	BASE	punct
124	7C	07/12		VERTICAL LINE	NON	BASE	punct
125	7D	07/13	}	RIGHT CURLY BRACKET	NON	BASE	punct
126	7E	07/14	~	TILDE	NON	BASE	
127	7F	07/15	DEL	DELETE	CTRL	NONDISP	
160	A0	10/00		NO-BREAK SPACE	NON	NONDISP	
161	A1	10/01	ก	KO KAI	CONS	BASE	cons alpha
162	A2	10/02	ข	KHO KHAI	CONS	BASE	cons alpha
163	A3	10/03	ช	KHO KHUT	CONS	BASE	cons alpha
164	A4	10/04	ค	KHO KHWAI	CONS	BASE	cons alpha
165	A5	10/05	ก	KHO KHON	CONS	BASE	cons alpha
166	A6	10/06	ฆ	KHO RAKHANG	CONS	BASE	cons alpha
167	A7	10/07	ง	NGO NGU	CONS	BASE	cons alpha
168	A8	10/08	จ	CHO CHAN	CONS	BASE	cons alpha
169	A9	10/09	ฉ	CHO CHING	CONS	BASE	cons alpha
170	AA	10/10	ช	CHO CHANG	CONS	BASE	cons alpha
171	AB	10/11	ซ	SO SO	CONS	BASE	cons alpha
172	AC	10/12	ฌ	CHO CHOE	CONS	BASE	cons alpha
173	AD	10/13	ญ	YO YING	CONS	BASE	cons alpha
174	AE	10/14	ฎ	DO CHADA	CONS	BASE	cons alpha
175	AF	10/15	ฏ	TO PATAK	CONS	BASE	cons alpha
176	B0	11/00	ฐ	THO THAN	CONS	BASE	cons alpha
177	B1	11/01	ฑ	THO NANGMONTHO	CONS	BASE	cons alpha
178	B2	11/02	ฒ	THO PHUTHAO	CONS	BASE	cons alpha
179	B3	11/03	ณ	NO NEN	CONS	BASE	cons alpha
180	B4	11/04	ด	DO DEK	CONS	BASE	cons alpha
181	B5	11/05	ต	TO TAO	CONS	BASE	cons alpha
182	B6	11/06	ถ	THO THUNG	CONS	BASE	cons alpha
183	B7	11/07	ท	THO THAHAN	CONS	BASE	cons alpha
184	B8	11/08	ธ	THO THONG	CONS	BASE	cons alpha
185	B9	11/09	น	NO NU	CONS	BASE	cons alpha
186	BA	11/10	บ	BO BAIMAI	CONS	BASE	cons alpha
187	BB	11/11	ป	PO PLA	CONS	BASE	cons alpha
188	BC	11/12	ผ	PHO PHUNG	CONS	BASE	cons alpha
189	BD	11/13	ฝ	FO FA	CONS	BASE	cons alpha

<i>Dec</i>	<i>Hex</i>	<i>Pos</i>	<i>Char</i>	<i>Name</i>	<i>Class</i>	<i>Level</i>	<i>Properties</i>
190	BE	11/14	พ	PHO PHAN	CONS	BASE	cons alpha
191	BF	11/15	ฟ	FO FAN	CONS	BASE	cons alpha
192	C0	12/00	ภ	PHO SAMPHAO	CONS	BASE	cons alpha
193	C1	12/01	ม	MOMA	CONS	BASE	cons alpha
194	C2	12/02	ย	YO YAK	CONS	BASE	cons alpha
195	C3	12/03	ร	RO RUA	CONS	BASE	cons alpha
196	C4	12/04	ฤ	RU	FV3	BASE	vowel alpha
197	C5	12/05	ล	LO LING	CONS	BASE	cons alpha
198	C6	12/06	ฬ	LU	FV3	BASE	vowel alpha
199	C7	12/07	ว	WO WAEN	CONS	BASE	cons alpha
200	C8	12/08	ศ	SO SALA	CONS	BASE	cons alpha
201	C9	12/09	ษ	SO RUSI	CONS	BASE	cons alpha
202	CA	12/10	ส	SO SUA	CONS	BASE	cons alpha
203	CB	12/11	ห	HO HIP	CONS	BASE	cons alpha
204	CC	12/12	ฬ	LO CHULA	CONS	BASE	cons alpha
205	CD	12/13	อ	O ANG	CONS	BASE	cons alpha
206	CE	12/14	ฮ	HO NOKHUK	CONS	BASE	cons alpha
207	CF	12/15	๑	PAIYANNOI	NON	BASE	punct
208	D0	13/00	๒	SARA A	FV1	BASE	vowel alpha
209	D1	13/01	๓	MAI HANฤๅAKAT	AV2	ABOVE	vowel alpha dead
210	D2	13/02	๔	SARA AA	FV1	BASE	vowel alpha
211	D3	13/03	๕	SARA AM	FV1	BASE	vowel alpha
212	D4	13/04	๖	SARA I	AV1	ABOVE	vowel alpha dead
213	D5	13/05	๗	SARA II	AV3	ABOVE	vowel alpha dead
214	D6	13/06	๘	SARA UE	AV2	ABOVE	vowel alpha dead
215	D7	13/07	๙	SARA UEE	AV3	ABOVE	vowel alpha dead
216	D8	13/08	๐	SARA U	BV1	BELOW	vowel alpha dead
217	D9	13/09	๑	SARA UU	BV2	BELOW	vowel alpha dead
218	DA	13/10	๒	PHINTHU	BD	BELOW	vowel alpha dead
219	DB	13/11		(This position shall not be used)	NON	NONDISP	
220	DC	13/12		THAI WORD BREAK	NON	NONDISP	punct
221	DD	13/13		(This position shall not be used)	NON	NONDISP	
222	DE	13/14		(This position shall not be used)	NON	NONDISP	
223	DF	13/15	฿	BAHT	NON	BASE	punct
224	E0	14/00	๓	SARA E	LV	BASE	vowel alpha

<i>Dec</i>	<i>Hex</i>	<i>Pos</i>	<i>Char</i>	<i>Name</i>	<i>Class</i>	<i>Level</i>	<i>Properties</i>
225	E1	14/01	๒	SARA AE	LV	BASE	vowel alpha
226	E2	14/02	๓	SARA O	LV	BASE	vowel alpha
227	E3	14/03	๔	SARA AI MAIMUAN	LV	BASE	vowel alpha
228	E4	14/04	๕	SARA AI MAIMALAI	LV	BASE	vowel alpha
229	E5	14/05	๖	LAKKHANGYAO	FV2	BASE	vowel alpha
230	E6	14/06	๗	MAIYAMOK	NON	BASE	punct
231	E7	14/07	๘	MAITAIKHU	AD2	TOP	alpha dead
232	E8	14/08	๙	MAI EK	TONE	TOP	tone alpha dead
233	E9	14/09	๐	MAI THO	TONE	TOP	tone alpha dead
234	EA	14/10	๑	MAI TRI	TONE	TOP	tone alpha dead
235	EB	14/11	๒	MAI CHATTAWA	TONE	TOP	tone alpha dead
236	EC	14/12	๓	THANTHAKHAT	AD1	TOP	alpha dead
237	ED	14/13	๔	NIKHAHIT	AD1	TOP	alpha dead
238	EE	14/14	๕	YAMAKKAN	AD3	ABOVE	alpha dead
239	EF	14/15	๖	FONGMAN	NON	BASE	punct
240	F0	15/00	๐	THAI DIGIT ZERO	NON	BASE	digit
241	F1	15/01	๑	THAI DIGIT ONE	NON	BASE	digit
242	F2	15/02	๒	THAI DIGIT TWO	NON	BASE	digit
243	F3	15/03	๓	THAI DIGIT THREE	NON	BASE	digit
244	F4	15/04	๔	THAI DIGIT FOUR	NON	BASE	digit
245	F5	15/05	๕	THAI DIGIT FIVE	NON	BASE	digit
246	F6	15/06	๖	THAI DIGIT SIX	NON	BASE	digit
247	F7	15/07	๗	THAI DIGIT SEVEN	NON	BASE	digit
248	F8	15/08	๘	THAI DIGIT EIGHT	NON	BASE	digit
249	F9	15/09	๙	THAI DIGIT NINE	NON	BASE	digit
250	FA	15/10	๐๑	ANGKHANKHU	NON	BASE	punct
251	FB	15/11	๐๒	KHOMUT	NON	BASE	punct
252	FC	15/12		(This position shall not be used)	NON	NONDISP	
253	FD	15/13		(This position shall not be used)	NON	NONDISP	
254	FE	15/14		(This position shall not be used)	NON	NONDISP	
255	FF	15/15		(This position shall not be used)	CTRL	NONDISP	

Table 3-3 Character classification table

Information about properties of characters in table 3-3 can also be used for functions to test the properties. Some of the functions are:

TACisalpha (c)	returns true if a character has alpha property
TACisalnum (c)	returns true if a character has alpha property or digit property
TACiscntrl (c)	returns true if a character is in CTRL class
TACisdigit (c)	returns true if a character has digit property
TACisgraph (c)	returns true if a character is not in NONDISP level excepts space
TACislower (c)	returns true if a character is a lowercase English alphabet (a-z)
TACisupper (c)	returns true if a character is an uppercase English alphabet (A-Z)
TACisprint (c)	returns true if a character in not in NONDISP level
TACisspace (c)	returns true if a character is space, formfeed, newline, carriage return, tab, vertical tab, and word break.
TACisxdigit (c)	returns true if a character is 0-9, A-F, a-f. (Thai digits are excluded)

4 General Rules about Displaying Thai Characters on Screen

When a Thai character is going to be output on the screen, there are two choices of display cell. The current character (also called following character) can either be displayed in the next display cell or composed with the previous character (also called leading character) and displayed in the previous display cell. The properties (chtype) of the following character (C_n) and the leading character (C_{n-1}) are used to determine which display cell the following character (C_n) should be in.

There are two rules to decide where the following character (C_n) should be displayed.

- If the following character (C_n) is a forward character, that is its chtype is NON, CONS, LV, FV1, FV2, or FV3, display it in the next display cell.
- If the following character (C_n) is a dead character, then examine the chtype of both the leading character and the following character. Check with figure 4-1 for the answer (N, C, or X).
 - If the answer is C (compose), then compose the following character (C_n) with the leading character (C_{n-1}) and display it in the previous display cell.
 - If the answer is N (next), then display the following character (C_n) in the next display cell.
 - If the answer is X (non-display), then treat it as usual control character.

<i>following</i> C_n		<i>FORWARD</i>						<i>DEAD</i>										<i>NOTES</i>		
		N	C	L	F	F	F	C	B	B	B	T	A	A	A	A	A		A	
		O	O	V	V	V	V	T	V	V	D	O	D	D	D	V	V		V	
<i>leading</i> C_{n-1}		N	S		1	2	3	R	1	2		N	1	2	3	1	2	3		
		1	2	3	4	5	6	0	7	8	9	10	11	12	13	14	15	16		
F O R W A R D	1	NON	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	NON COMPOSIBLES
	2	CONS	N	N	N	N	N	N	X	C	C	C	C	C	C	C	C	C	C	CONSONANTS
	3	LV	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ł (0xE0) łł (0xE1) Ł (0xE2) Ł (0xE3) Ł (0xE4)
	4	FV1	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ż (0xD0) ʀ (0xD2) ʀ (0xD3)
	5	FV2	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ʀ (0xE5)
	6	FV3	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ʀ (0xC4) ʀ (0xC6)
D E A D	0	CTRL	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	
	7	BV1	N	N	N	N	N	N	X	N	N	N	C	C	N	N	N	N	N	ˆ (0xD8)
	8	BV2	N	N	N	N	N	N	X	N	N	N	C	N	N	N	N	N	N	ˆ (0xD9)
	9	BD	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ˆ (0xDA)
	10	TONE	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ˆ (0xE8) ˆ (0xE9) ˆ (0xEA) ˆ (0xEB)
	11	AD1	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ˆ (0xED) ˆ (0xEC)
	12	AD2	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ˆ (0xE7)
	13	AD3	N	N	N	N	N	N	X	N	N	N	N	N	N	N	N	N	N	ˆ (0xEE)
	14	AV1	N	N	N	N	N	N	X	N	N	N	C	C	N	N	N	N	N	ˆ (0xD4)
	15	AV2	N	N	N	N	N	N	X	N	N	N	C	N	N	N	N	N	N	ˆ (0xD1) ˆ (0xD6)
16	AV3	N	N	N	N	N	N	X	N	N	N	C	N	C	N	N	N	N	ˆ (0xD5) ˆ (0xD7)	

N = display the following character in the next display cell
C = compose the following character with the leading character and display in the same display cell as the leading character
X = non-display

Figure 4 -1 Basic output method of charcters by considering chtypes of both leading and following characters

5 *Input Method*

5.1 *Basic rules*

Input method for Thai characters is direct mapping method like English. No matter a Thai character is a forward character or a dead character, it is entered character by character and displayed from left to right. There are some basic rules concerning the input sequence.

- Every display cell must begin with character in base level. If the character is in CONS class, it can be composed with above vowel, below vowel, above diacritic, below diacritic or tone mark (character in AV, BV, AD, BD or TONE class).
- Character in CONS class may be followed by above vowel or below vowel (character in AV1-AV3 or BV1-BV2 class).
- Character in CONS class (including those followed by above or below vowel) may be followed by tone mark, below diacritic or above diacritic (character in TONE, BD, AD1 or AD2 class).

The above rules are used to construct the syntax diagram (figure 5-1). Input sequence that does not conform the syntax diagram is certainly not a proper Thai word or sentence, thus will be rejected. However, the input sequence of any word or sentence that is meaningless but conform to the syntax diagram is accepted.

In a typical case, the word กั for example, consists of 3 characters, one in base level, one in above level and one in top level. The syntax diagram will only accept the correct input sequence กั (0xB7 0xD5 0xE8), it will reject the illegal input sequence กั (0xB7 0xE8 0xD5).

Since input routine works independent of output routine, the input routine should block further input if an illegal input sequence is detected, then it should interrupt and notify the user by giving a warning bell or some other way of warning.

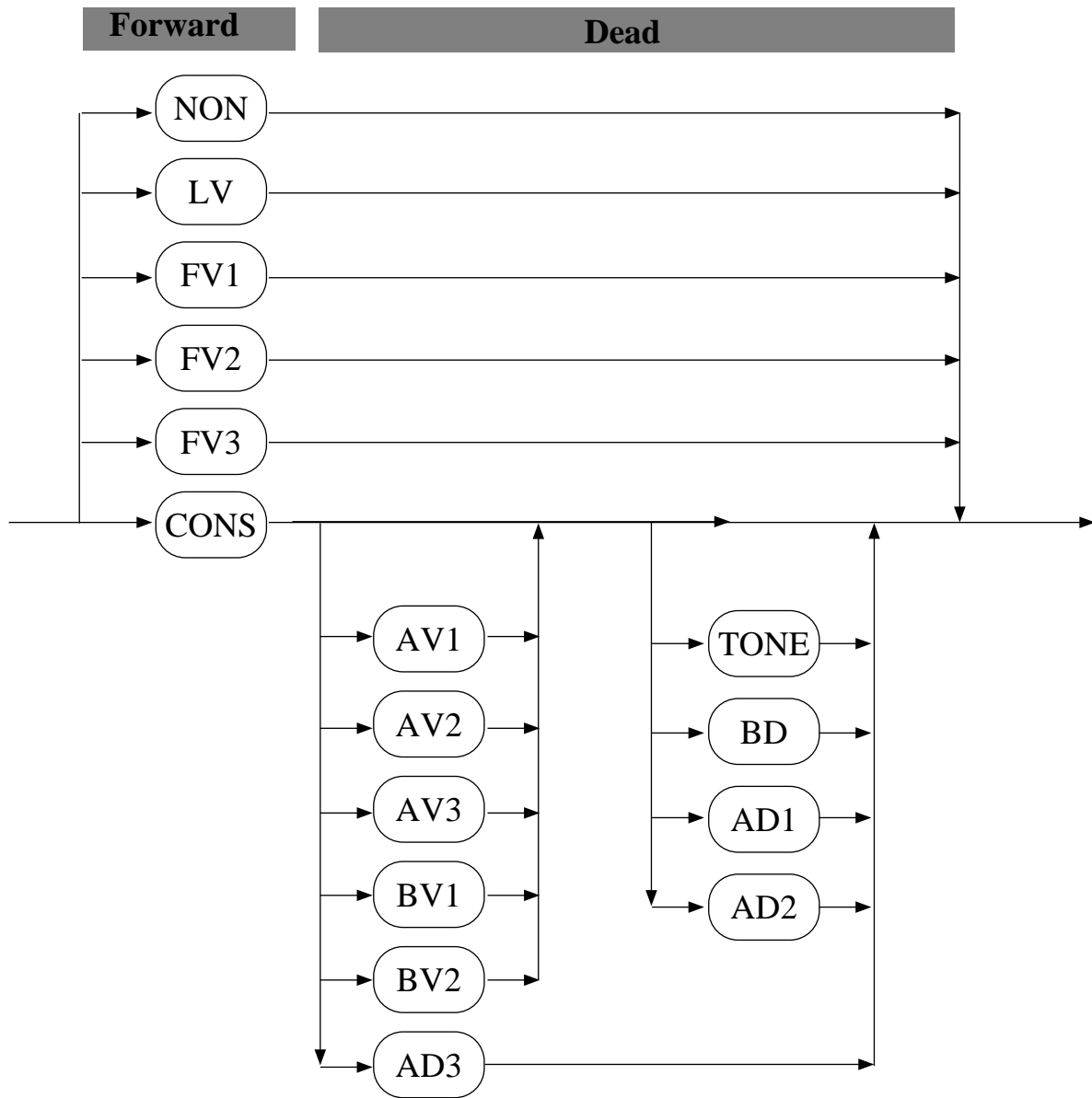


Figure 5 -1 Syntax diagram showing proper input sequence of character strings

5.2 Input Sequence Checking

An API function (`TACcomposable`) has been defined for checking of input sequence. There are three modes of checking.

- **Mode 0**

This is the *passthrough* mode. It allows the application program to handle the checking of input sequence. The application program will be responsible to

call the proper input routine to get input data for further processing. When the program terminates, the mode should be set back to mode 1 (default).

- **Mode 1 (default)**

This is the *basic check* mode. It obeys the rules following the syntax diagram in the Basic Rules section. Application program may add additional checking if needed. This is the default mode to Thai system.

- **Mode 2**

This is the *strict* mode. This mode is the default mode with addition of more conditions in order to reject obvious illegal input sequence (illegal composibility) as defined figure 5-2 and 5-3 . The strict mode conforms to the new syntax diagram (figure 5-4). It improves the correctness of input sequence.

The following API function can be used to check the composibility of leading and following characters.

TACcomposable (chn, chn1)

chn = following character, chn1 = leading character

The function returns char which has value "A", "C", "S", "R", or "X" as stated in figure 5-2 and 5-3.

Value	Meaning	Accept/Reject			Where to display if accepted
		Mode 0	Mode 1	Mode 2	
A	accept	accept	accept	accept	next display cell
C	accept (compose)	accept	accept	accept	same display cell as leading character
S	strict mode reject	accept	accept	reject	next display cell
R	reject	accept	reject	reject	next display cell
X	accept (non-display)	accept	accept	accept	non-display

Figure 5 -2 Meaning of values returned by input sequence checking routine in figure 5-3.

<i>following</i> C_n		<i>FORWARD</i>							<i>DEAD</i>									<i>NOTES</i>		
		N O N	C O N S	L V	F V 1	F V 2	F V 3	C T R L	B V 1	B V 2	B D	T O N E	A D 1	A D 2	A D 3	A V 1	A V 2		A V 3	
		1	2	3	4	5	6	0	7	8	9	10	11	12	13	14	15		16	
<i>F O R W A R D</i>	1	NON	A	A	A	S	S	A	X	R	R	R	R	R	R	R	R	R	R	NON COMPOSIBLES
	2	CONS	A	A	A	A	S	A	X	C	C	C	C	C	C	C	C	C	C	CONSONANTS
	3	LV	S	A	S	S	S	S	X	R	R	R	R	R	R	R	R	R	R	ł (0xE0) ʟ (0xE1) Ł (0xE2) ʟ̇ (0xE3) ł̇ (0xE4)
	4	FV1	S	A	S	A	S	A	X	R	R	R	R	R	R	R	R	R	R	ʒ (0xD0) ɹ (0xD2) ɹ̇ (0xD3)
	5	FV2	A	A	A	A	S	A	X	R	R	R	R	R	R	R	R	R	R	ɹ (0xE5)
	6	FV3	A	A	A	S	A	S	X	R	R	R	R	R	R	R	R	R	R	ɸ (0xC4) ɸ̇ (0xC6)
<i>D E A D</i>	0	CTRL	A	A	A	A	A	A	X	R	R	R	R	R	R	R	R	R	R	
	7	BV1	A	A	A	A	S	A	X	R	R	R	C	C	R	R	R	R	R	ɹ̇ (0xD8)
	8	BV2	A	A	A	S	S	A	X	R	R	R	C	R	R	R	R	R	R	ɹ̇ (0xD9)
	9	BD	A	A	A	S	S	A	X	R	R	R	R	R	R	R	R	R	R	˙ (0xDA)
	10	TONE	A	A	A	A	A	A	X	R	R	R	R	R	R	R	R	R	R	ˆ (0xE8) ˆ̇ (0xE9) ˆ̇ (0xEA) ˆ̇+ (0xEB)
	11	AD1	A	A	A	S	S	A	X	R	R	R	R	R	R	R	R	R	R	ˆ̇ (0xED) ˆ̇ (0xEC)
	12	AD2	A	A	A	S	S	A	X	R	R	R	R	R	R	R	R	R	R	ˆ̇ (0xE7)
	13	AD3	A	A	A	S	S	A	X	R	R	R	R	R	R	R	R	R	R	ˆ̇ (0xEE)
	14	AV1	A	A	A	S	S	A	X	R	R	R	C	C	R	R	R	R	R	ˆ (0xD4)
	15	AV2	A	A	A	S	S	A	X	R	R	R	C	R	R	R	R	R	R	ˆ̇ (0xD1) ˆ̇ (0xD6)
16	AV3	A	A	A	S	S	A	X	R	R	R	C	R	C	R	R	R	R	ˆ̇ (0xD5) ˆ̇ (0xD7)	

Figure 5 -3 Map showing value returned by input sequence checking routine which considers chtypes of both leading and following characters.

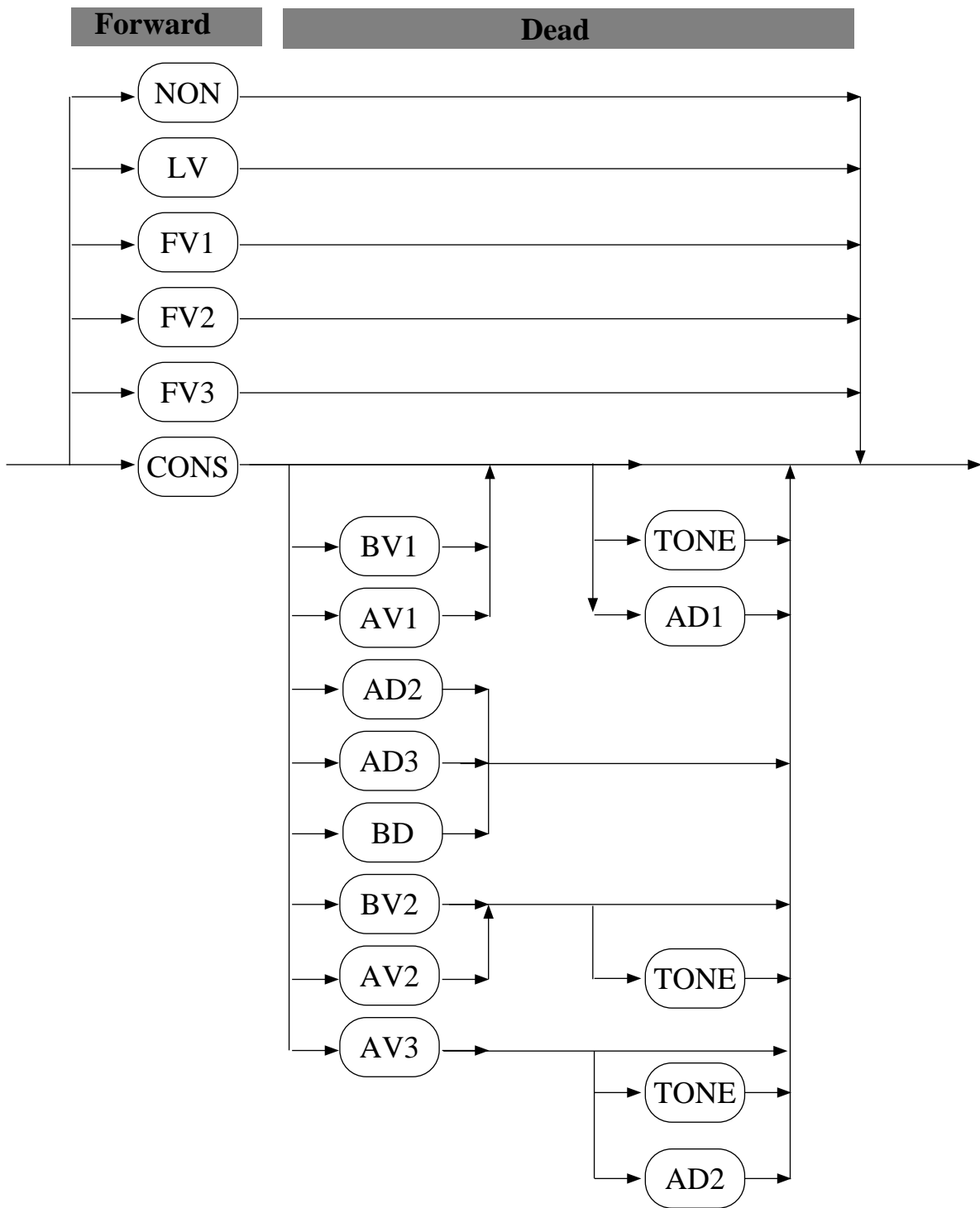


Figure 5 -4 Syntax diagram showing input sequence of character strings in strict mode

6 Output Method

6.1 Character Strings

Character strings to be output on screen must use TACTIS code, the character code defined in the basic character set. Hence, character strings with other kind of character codes have to be converted to TACTIS first. TACTIS code shall be used as standard code for exchange of information in this standard.

6.2 Combining above/below level character with base level character

Characters that are classified as dead characters (AV, BV, AD, BD, or TONE class character) can be displayed in the same display cell as consonant (CONS class character) by having the input sequence conformed to syntax diagram (figure 5-4). Below is a summary from the study of possible ways to combine dead character and consonant.

- The maximum number of characters in a display cell is 3. Therefore, there can be 1, 2 or 3 characters in a display cell.
- Each display cell must have a character that is in NON, CONS, LV, or FV class. If the display cell contains a consonant (CONS class character), the consonant may be composed with 0-2 dead characters and display in the same display cell.
- There are 51 legal combinations (compositions) of dead characters which is shown in figure 6-2. Any Thai system should have at least 47 out of 51 legal combinations of dead characters. The four optional combinations are (0xBE), (0xDE), (0xE7) and (0xF7). Thai system to be developed in the future should support all 51 legal combinations.

6.3 Data Structure of Display Cell

A data structure for display cell called THAI_CELL has been defined for use in information exchange between screen display and application program. THAI_CELL consists of 4 bytes namely mid, hilo, attr and style bytes.

mid byte = character code of base level character

hilo byte = character code of 0-2 (combining) dead characters

	<i>Bit</i>							
	7	6	5	4	3	2	1	0
Value	1							

<i>Bit</i>				<i>Char</i>
6	5	4	3	
0	0	0	0	Reserved for future use
0	0	0	1	Reserved for future use
0	0	1	0	Reserved for future use
0	0	1	1	Reserved for future use
0	1	0	0	Reserved for future use
0	1	0	1	Reserved for future use
0	1	1	0	space
0	1	1	1	◌̣ (0xD8)
1	0	0	0	◌̣̣ (0xD9)
1	0	0	1	◌̣̣̣ (0xDA)
1	0	1	0	◌̣̣̣̣ (0xD1)
1	0	1	1	◌̣̣̣̣̣ (0xD4)
1	1	0	0	◌̣̣̣̣̣̣ (0xD5)
1	1	0	1	◌̣̣̣̣̣̣̣ (0xD6)
1	1	1	0	◌̣̣̣̣̣̣̣̣ (0xD7)
1	1	1	1	◌̣̣̣̣̣̣̣̣̣ (0xEE)

<i>Bit</i>			<i>Char</i>
2	1	0	
0	0	0	space
0	0	1	◌̣̣̣ (0xE8)
0	1	0	◌̣̣̣̣ (0xE9)
0	1	1	◌̣̣̣̣̣ (0xEA)
1	0	0	◌̣̣̣̣̣̣ (0xEB)
1	0	1	◌̣̣̣̣̣̣̣ (0xEC)
1	1	0	◌̣̣̣̣̣̣̣̣ (0xED)
1	1	1	◌̣̣̣̣̣̣̣̣̣ (0xE7)

Bit 2-0, tone marks and diacritics

Bit 6-3, non-spacing vowels and YAMMAKKAN

Note: The eight bit (Bit 7) is always set to 1.

Figure 6 -1 Hilo byte

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0												space		๐	๑	๒
1												๓	๔	๕	๖	๗
2												๘	๙	๐	๑	๒
3												๓	๔	๕	๖	๗
4												๘	๙	๐	๑	๒
5												๓				
6												๔				
7												๕			๐๑	๐๒
8												๖		๐	๑	๒
9												๗		๐	๑	
A												๘		๐	๑	
B												๙		๐	๑	
C												๐		๐	๑	
D												๑		๐		
E												๒		๐		
F																

□ Optional ■ Invalid

Figure 6 -2 Character codes for fifty-one valid compositions of dead characters

The data structure THAI_CELL is useful for application programmers who develop overlapping window, screen capture, or remote screen control applications. However, the 2-byte data structure that consists of mid and hilo bytes or the 4-byte THAI_CELL are not a standard for typical information exchange.

Some examples of API functions that will make use of THAI_CELL are

```
TACvreadscr (*buf, x, y, w, h);
TACvgetchar (thai_cell);
TACvloadat (id, x, y);
TACvloadfrom (id);
TACvsaveto (mode, id);
```

6.4 Displaying character strings that have illegal input sequence

It is possible to have character strings entered improperly on a Thai system, transferred to another Thai system and displayed. Therefore, it is important to have same method to display character string that contains illegal character sequence.

An *illegal character sequence* means the character string contains a pair of leading-following characters that has composibility value of "R" or "S" according to figure 5-2 and 5-3. In this case, the following character must be displayed in the next character cell. By complying to this, all applications will display character strings with illegal character sequence alike. One benefit of following the guideline is that it helps user spot the mistake easily.

There are two ways of displaying illegal character sequence. One way is to display illegal character sequence on base level. Another way is to display illegal character sequence on top,above or below level correspondingly.

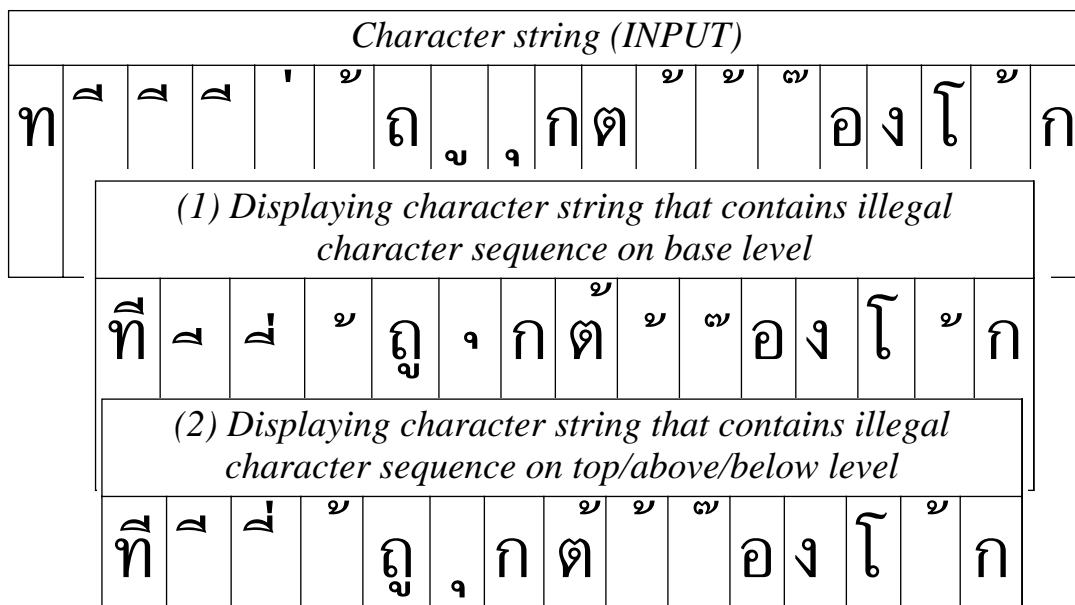


Figure 6 -3 Output a character string according to rules in figure 5-2 and 5-3

7 Interpreting Control Characters

How control characters should be interpreted when the screen is in teletype (TTY) mode is described in table 7-1 below.

<i>Control code</i>	<i>What it does</i>
BEL (bell)	No change to the screen. Short beep sound.
BS (backspace)	Move the cursor one position to the left. If the cursor is at the first column, it will stop moving. If the cursor is at a display cell which contains Thai characters in top/above/below level, the cursor will have to compensate for all extra characters in the display cell before moving to next display cell. For example, if there are a consonant, a below vowel and a tone mark in a display cell, user will have to hit BS three times to pass to the next display cell. Note that BS is non destructive backspace.
HT (horizontal tab)	Fill spaces until the cursor reaches next tab stop position. By standard, tab stops every 8 columns. The left most column is 0, tab stops at column 8, 16, 24, 32, ...
LF (line feed)	Move cursor down by one line. If the cursor is at the last line of the screen, move the screen up by one line, the cursor remains at the last line.
FF (form feed)	Clear the screen and move the cursor to the upper left most position.
CR (carriage return)	Move the cursor to the first column
other control characters	Ignore. No change to the screen.

Table 7 -1 Interpretation of control character in teletype (TTY) mode

8 Text Editing

Application programs that have text editing feature should follow the following guidelines.

- Cursor movement operation moves from one display cell to another.
- Delete character operation deletes one character each time.
- Delete cell operation deletes one display cell each time.

<i>Function</i>	<i>What it does</i>
forward cell	Move the cursor one display cell to the right
backward cell	Move the cursor one display cell to the left
delete current cell *(delete next cell)	Delete all characters in the display cell that the cursor resides. (There could be 1-3 characters in a Thai display cell.)
delete previous char	Delete a character in the display cell to the left of the cursor.

*Meaning delete next cell if the cursor is a vertical bar insertion point.

Table 8 -1 Meaning of cursor movement and deletion in text editing

<i>Keytop</i>	<i>Function</i>
⇒	forward cell
⇐	backward cell
delete	delete current cell
backspace	delete previous cell

Table 8 -2 Keytops and their functions in text editing

9 Appendix

9.1 The File: TIS.H

```
/* NECTEC THAI SOFTWARE STANDARD PROJECT 1989-1991
FILE TIS.H
This include file defines the names and character codes of
(1) Standard ASCII control codes,
(2) Thai Industrial Standard characters (ASCII TIS 620-2533),
(3) Dot Matrix Printer Code Extension (TSI 988-2533),
(4) WTT2.0 standard internal code extension for Hilo Byte
encoding
The names are written as defined in the future announcement
of ISO-/ECMA 8859-?? for Thailand.

WRITTEN BY: THAWEEESAK KOANANTAKOOL
IPIED, THAMMASAT UNIVERSITY
LAST UPDATE: 18 FEBRUARY 1991
*/

/* PART 1 - STANDARD ASCII CODES (ISO-646) */

#define NULL          0
#define NUL          0
#define SOH          1      /* 0x01 START OF HEADING */
#define STX          2      /* 0x02 START OF TEXT */
#define ETX          3      /* 0x03 END OF TEXT */
#define EOT          4      /* 0x04 END OF TRANSMISSION */
#define ENQ          5      /* 0x05 ENQUIRE */
#define ACK          6      /* 0x06 ACKNOWLEDGE */
#define BEL          7      /* 0x07 BELL */
#define BS           8      /* 0x08 BACKSPACE */
#define HT           9      /* 0x09 HORIZONTAL TAB */
#define LF           10     /* 0x0A LINE FEED */
#define VT           11     /* 0x0B VERTICAL TAB */
#define FF           12     /* 0x0C FORM FEED */
#define CR           13     /* 0x0D CARRIAGE RETURN */
#define SO           14     /* 0x0E SHIFT OUT */
#define SI           15     /* 0x0F SHIFT IN */
#define DLE          16     /* 0x10 DATA LINK ESCAPE */
#define DC1          17     /* 0x11 DEVICE CONTROL 1 */
#define XON          DC1    /* 0x11 X-ON */
#define DC2          18     /* 0x12 DEVICE CONTROL 2 */
#define DC3          19     /* 0x13 DEVICE CONTROL 3 */
#define XOFF         DC3    /* 0x13 X-OFF */
#define DC4          20     /* 0x14 DEVICE CONTROL 4 */
#define NAK          21     /* 0x15 NEGATIVE ACKNOWLEDGE */
#define SYN          22     /* 0x16 SYNCHRONOUS IDLE */
#define ETB          23     /* 0x17 END OF TRANSMISSION BLOCK */
#define CAN          24     /* 0x18 CANCEL */
```

```

#define EM          25      /* 0x19 END OF MEDIUM */
#define SUB         26      /* 0x1A SUBSTITUTE CHARACTER */
#define ESC         27      /* 0x1B ESCAPE */
#define FS          28      /* 0x1C FILE SEPARATOR */
#define GS          29      /* 0x1D GROUP SEPARATOR */
#define RS          30      /* 0x1E RECORD SEPARATOR */
#define US          31      /* 0x1F UNIT SEPARATOR */
#define IS4         28      /* 0x1C INFORMATION SEPARATOR 4 */
#define IS3         29      /* 0x1D INFORMATION SEPARATOR 3 */
#define IS2         30      /* 0x1E INFORMATION SEPARATOR 2 */
#define IS1         31      /* 0x1F INFORMATION SEPARATOR 1 */
#define SP          32      /* 0x20 SPACE */
#define DEL         127     /* 0x7F DELETE */

```

```

/* PART 2 - THAI INDUSTRIAL STANDARD -- TIS 620-2529(1986)
NAMING CONVENTION AS TISO PROPOSED TO ISO/ECMA FOR
DRAFT ISO-8859-xx FOR THAILAND.
UPDATED 18 FEB 91 ACCORDING TO TC536 RESOLUTION
*/

```

```

#define KOKAI       161     /* ก 0xA1 */
#define KHOKHAI     162     /* ข 0xA2 */
#define KHOKHUAT    163     /* ค 0xA3 */
#define KHOKHWAI    164     /* ฉ 0xA4 */
#define KHOKHON     165     /* ช 0xA5 */
#define KHORAKHANG  166     /* ฆ 0xA6 */
#define NGONGU      167     /* ง 0xA7 */
#define CHOCHAN     168     /* จ 0xA8 */
#define CHOCHING    169     /* ฉ 0xA9 */
#define CHOCHANG    170     /* ฌ 0xAA */
#define SOSO        171     /* ซ 0xAB */
#define CHOCHOE     172     /* ฌ 0xAC */
#define YOYING      173     /* ญ 0xAD */
#define DOCHADA     174     /* ฎ 0xAE */
#define TOPATAK     175     /* ฏ 0xAF */
#define THOTHAN     176     /* ฐ 0xB0 */
#define THONANGMONT 177     /* ท 0xB1 */
#define THOPHUTHAO  178     /* ฒ 0xB2 */
#define NONEN       179     /* ณ 0xB3 */
#define DODEK       180     /* ด 0xB4 */
#define TOTAO       181     /* ต 0xB5 */
#define THOTHUNG    182     /* ถ 0xB6 */

```

```

#define THOTHAHAN      183    /* ท 0xB7 */
#define THOTHONG       184    /* ธ 0xB8 */
#define NONU           185    /* น 0xB9 */
#define BOBAIMAI       186    /* บ 0xBA */
#define POPLA          187    /* ป 0xBB */
#define PHOPUENG       188    /* ผ 0xBC */
#define FOFA           189    /* ฝ 0xBD */
#define PHOPHAN        190    /* พ 0xBE */
#define FOFAN          191    /* ฟ 0xBF */
#define PHOSAMPHAO    192    /* ภ 0xC0 */
#define MOMA           193    /* ม 0xC1 */
#define YOYAK          194    /* ย 0xC2 */
#define RORUA         195    /* ร 0xC3 */
#define RU             196    /* ฤ 0xC4 */
#define LOLING        197    /* ล 0xC5 */
#define LU             198    /* ฌ 0xC6 */
#define WOWAEN        199    /* ฌ 0xC7 */
#define SOSALA        200    /* ศ 0xC8 */
#define SORUSI        201    /* ส 0xC9 */
#define SOSUA         202    /* ส 0xCA */
#define HOHIP         203    /* ห 0xCB */
#define LOCHULA       204    /* พ 0xCC */
#define OANG          205    /* อ 0xCD */
#define HONOKHUK      206    /* ฮ 0xCE */
#define PAIYANNOI    207    /* ฦ 0xCF */
#define SARA_A        208    /* ๐ 0xD0 */
#define MAIHUNAKAT    209    /* ๑ 0xD1 */
#define SARA_AA       210    /* ๒ 0xD2 */
#define SARA_AM       211    /* ๓ 0xD3 */
#define SARA_I        212    /* ๔ 0xD4 */
#define SARA_II       213    /* ๕ 0xD5 */
#define SARA_UE       214    /* ๖ 0xD6 */
#define SARA_UEE      215    /* ๗ 0xD7 */
#define SARA_U        216    /* ๘ 0xD8 */
#define SARA_UU       217    /* ๙ 0xD9 */

```

```

#define PHINTHU          218    /*  . 0xDA */
#define BAHT             223    /*  ฿ 0xDF */
#define SARA_E           224    /*  ใ 0xE0 */
#define SARA_AE          225    /*  ใ 0xE1 */
#define SARA_O           226    /*  ใ 0xE2 */
#define MAIMUAN          227    /*  ใ 0xE3 */
#define MAIMALAI         228    /*  ใ 0xE4 */
#define LAKKHANGYAO      229    /*  ใ 0xE5 */
#define MAIYAMOK         230    /*  ใ 0xE6 */
#define MAITAIKHU        231    /*  ใ 0xE7 */
#define MAIEK            232    /*  ใ 0xE8 */
#define MAITHO           233    /*  ใ 0xE9 */
#define MAITRI           234    /*  ใ 0xEA */
#define MAICHATTAWA      235    /*  ใ 0xEB */
#define THANTHAKHAT      236    /*  ใ 0xEC */
#define NIKHAHIT         237    /*  ใ 0xED */
#define YAMAKKAN         238    /*  ใ 0xEE */
#define FONGMAN          239    /*  ใ 0xEF */
#define THAIZERO         240    /*  ใ 0xF0 */
#define THAIONE          241    /*  ใ 0xF1 */
#define THAITWO          242    /*  ใ 0xF2 */
#define THAITHREE        243    /*  ใ 0xF3 */
#define THAIFOUR         244    /*  ใ 0xF4 */
#define THAIFIVE         245    /*  ใ 0xF5 */
#define THAISIX          246    /*  ใ 0xF6 */
#define THAISEVEN        247    /*  ใ 0xF7 */
#define THAIEIGHT        248    /*  ใ 0xF8 */
#define THAININE         249    /*  ใ 0xF9 */
#define ANGKHANKHU       250    /*  ใ 0xFA */
#define KHOMUT           251    /*  ใ 0xFB */

/* PART 3 -THAI INDUSTRIAL STANDARD EXTENDED FOR PRINTER */
/* TIS 988-2533 */

#define MAIHUNAKAT_EK    0x80    /* TIS 988-2533 */
#define SARA_I_EK        0x84    /* TIS 988-2533 */
#define SARA_I_THANTHAKHAT 0x88    /* TIS 988-2533 */
#define SARA_II_EK       0x89    /* TIS 988-2533 */

```

```

#define SARA_UE_EK          0x8D    /* TIS 988-2533 */
#define SARA_UEE_EK        0x91    /* TIS 988-2533 */

#define PR_UM_EK           0x95    /* TIS 988-2533 */
#define PR_BOTTOM_O       0xDB    /* TIS 988-2533 */
#define PR_TOP_O          0xDC    /* TIS 988-2533 */
#define PR_TOP_MAIMUAN    0xDD    /* TIS 988-2533 */
#define PR_TOP_MAIMALAI   0xDE    /* TIS 988-2533 */

```

```

/* PART 4 - WTT 2.0 Hilo Byte Encoding */
/* Revision 3 Certified by TAPIC on March 27 1991 */

```

```

enum HiloByte {
    HSPLCE=0xB0, HEK, HTHO, HTRI, HCHATTAWA, HTHANTHAKHAT, HNIKHAHIT, HTAIKHU,
    HU=0xB8, HU_EK, HU_THO, HU_TRI, HU_CHATTAWA, HU_THANTHAKHAT, HU_NUKHAHIT,
    HUU=0xC0, HUU_EK, HUU_THO, HUU_TRI, HUU_CHATTAWA,
    HPHINTHU=0xC8,
    HHUNAKAT=0xD0, HHUNAKAT_EK, HHUNAKAT_THO, HHUNAKAT_TRI,
    HHUNAKAT_CHATTAWA,
    HI=0xD8, HI_EK, HI_THO, HI_TRI, HI_CHATTAWA,
    HI_THANTHAKHAT, HI_NIKHAHIT,
    HII=0xE0, HII_EK, HII_THO, HII_TRI, HII_CHATTAWA,
    HII_TAIKHU=0xE7,
    HUE=0xE8, HUE_EK, HUE_THO, HUE_TRI, HUE_CHATTAWA,
    HUEE=0xF0, HUEE_EK, HUEE_THO, HUEE_TRI, HUEE_CHATTAWA,
    HUEE_TAIKHU=0xF7,
    HYANMAKKAN=0xF8
};

```

```

/* Allow use of the symbol KARAN in place of THANTHAKHAT */
#define KARAN                THANTHAKHAT
#define HKARAN               HTHAKANTHAKHAT
#define HU_KARAN             HU_THAKANTHAKHAT
#define HI_KARAN             HI_THAKANTHAKHAT

```

```

/*----- END OF TIS.H -----*/

```